



## DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

<b>(51) Classification internationale des brevets <sup>6</sup> :</b>  <b>C12Q 1/68</b>	<b>A1</b>	<b>(11) Numéro de publication internationale:</b> <b>WO 97/36003</b>  <b>(43) Date de publication internationale:</b> 2 octobre 1997 (02.10.97)
<b>(21) Numéro de la demande internationale:</b> PCT/FR97/00518  <b>(22) Date de dépôt international:</b> 24 mars 1997 (24.03.97)  <b>(30) Données relatives à la priorité:</b> 96/03835      22 mars 1996 (22.03.96) <b>FR</b>  <b>(71) Déposant (pour tous les Etats désignés sauf US):</b> BIO MERIEUX [FR/FR]; 58, avenue Leclerc, F-69280 Marcy- l'Etoile (FR).  <b>(72) Inventeurs; et</b> <b>(75) Inventeurs/Déposants (US seulement):</b> CARLOTTI, Amaud [FR/FR]; 9, passage Feuillat, F-69003 Lyon (FR). VIL- LARD, Jean [FR/FR]; 22, rue Léo-Trouilhet, F-69008 Lyon (FR).  <b>(74) Mandataire:</b> ROPITAL-BONVARLET, Claude; Cabinet Beau de Loménie, 51, avenue Jean-Jaurès, Boîte postale 7073, F- 69301 Lyon Cédex 07 (FR).		<b>(81) Etats désignés:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, brevet ARIPO (GH, KE, LS, MW, SD, SZ, UG), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Publiée</b>  <i>Avec rapport de recherche internationale.</i> <i>Avant l'expiration du délai prévu pour la modification des</i> <i>revendications, sera republiée si de telles modifications sont</i> <i>reçues.</i>
<b>(54) Title:</b> PROBE, NUCLEIC ACID SEQUENCES, METHOD PARTICULARLY SUITABLE FOR DETECTING AND IDENTIFYING YEASTS, AND GENETIC TOOLS THEREFOR		
<b>(54) Titre:</b> SONDE, SEQUENCE NUCLEIQUES ET PROCEDE NOTAMMENT, DE DETECTION ET D'IDENTIFICATION DE LEVURES, DE MEME QUE LES OUTILS GENETIQUES CORRESPONDANTS		
<b>(57) Abstract</b>  The detection and characterisation of yeasts, particularly those of medical and industrial interest, is disclosed. The aim is to provide powerful, reliable and simply and reproducibly usable DNA probes specific for numerous yeasts. For this purpose, among others, a yeast-specific and/or -intraspecific detection probe is provided which consists of at least one part of at least one DNA and/or RNA fragment F from rDNA that does not code for rRNA synthesis and is located in the IGR (rDNA 25S and rDNA 18S). This iterative rDNA includes at least two repeated subsequences with a mutual homology of $\geq 50\%$ . A detection method using said probe, as well as nucleic acid sequences capable of being included in the probe, are also disclosed. Said probe is useful for detecting/identifying yeast (e.g. <i>C. krusei</i> , <i>G. candidum</i> ), as an epidemiological marker or strain tracer, and in anti-yeast therapy.		
<b>(57) Abrégé</b>  Le domaine de l'invention est celui de la détection et la caractérisation de levures notamment celles d'intérêt médical et industriel. Le but visé par l'invention est la fourniture de sondes nucléiques spécifiques de nombreuses levures, et par ailleurs performantes, fiables et utilisables de manière simple et reproductible. Pour satisfaire à ce but parmi d'autres, l'invention propose ainsi une sonde de détection spécifique et/ou intraspécifique de levure, caractérisée en ce qu'elle est constituée par au moins une partie d'au moins un fragment F d'ADN et/ou d'ARN, provenant d'ADNr qui ne code pas pour la synthèse d'ARNr et qui est localisé dans l'IGR (ADNr 25S et ADNr 18S). Cet ADNr est de type itératif comprend au moins deux sous-séquences répétées, avec une homologie entre sous-séquences $\geq 50\%$ . L'invention concerne en outre un procédé de détection mettant en oeuvre cette sonde, ainsi que des séquences nucléiques susceptibles d'être incluses dans la sonde. Application: détection/identification de levure ( <i>C.krusei</i> , <i>G.candidum</i> ) par exemple, marqueur épidémiologique, traceur de souches, thérapeutique anti-levures.		

### UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Bésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

SONDE, SEQUENCE NUCLEIQUES ET PROCEDE NOTAMMENT, DE DETECTION ET D'IDENTIFICATION DE LEVURES, DE MEME QUE LES OUTILS GENETIQUES CORRESPONDANTS

5     **DOMAINE TECHNIQUE :**

La présente invention concerne la détection et la caractérisation de levures. En particulier elle fournit des sondes spécifiques, notamment nucléiques, et un modèle de développement des dites sondes pour la détection et, en particulier, l'identification et  
10 le typage (caractérisation infra-spécifique) des souches de levures de différentes espèces d'intérêt médical et industriel.

**ART ANTERIEUR :**

15 Les levures sont des champignons, micro-organismes eucaryotes, chez lesquels la forme unicellulaire est prédominante (Barnett J., W., R., W. Payne and D. Yarrow., "Yeasts characteristics and identification, 2nd Edition, Cambridge university press, Cambridge, 1991). Elles sont divisées en genres (e.g. *Candida*) et espèces (e.g. *Candida krusei*) sur la base de critères de reproduction sexuée (lorsqu'elle existe),  
20 morphologiques, physiologiques, chimiotaxonomiques et moléculaires.

Les levures sont ubiquitaires. Généralement leur innocuité est bien établie. Elles sont souvent utilisées comme agent de transformation (auxiliaires technologiques) dans les procédés agro-alimentaires, mais certaines sont parfois pathogènes. Les levures sont généralement considérées comme étant des pathogènes opportunistes.

25 Les levures pathogènes opportunistes intéressent le domaine de la mycologie médicale. Ces levures sont responsables d'infection superficielles, cutanées, sous-cutanées, et systémiques. Ces dernières sont les plus sérieuses en raison des taux de mortalité élevés qui leur sont imputés. Les levures pathogènes appartiennent le plus souvent au genre *Candida*. Parmi ce genre les espèces plus particulièrement  
30 importantes sont *Candida albicans*, *Candida krusei*, *Candida tropicalis*, *Candida lusitanae*, *Candida (Torulopsis) glabrata*. Les infections qu'elles provoquent sont appelées "Candidoses".

La prévalence et la pathogénécité de ces levures opportunistes sont élevées chez les sujets immunodéprimés, notamment chez les patients neutropéniques, atteints  
35 de SIDA, cancéreux, etc...

Toute la difficulté du traitement de ce type d'infections, provient du fait que certaines espèces sont résistantes à certains antifongiques, notamment au fluconazole (e.g. *C. krusei*, *C. glabrata*). En outre, pour une même espèce certaines souches présentent un caractère de résistance alors que d'autres sont sensibles aux antifongiques considérés (e.g. *C. albicans*, *C. lusitaniae*).

En conséquence, une détection rapide et une identification précise et fiable ainsi qu'un typage des souches de ces espèces sont de plus en plus nécessaires, notamment pour permettre la mise en place d'un traitement antifongique approprié. Il existe, en outre, un réel besoin en marqueurs épidémiologiques fiables, à des fins d'études épidémiologiques dans le domaine médical, ou de traçage des souches dans les industries agroalimentaires ou de fermentation.

Mais force est de constater que, les méthodes classiques d'identification des levures ne donnent pas entière satisfaction à tous ces égards.

On connaît ainsi notamment celles décrites par Barnett J., W., R. W. Payne and D. Yarrow., dans "Yeasts characteristics and identification", 2nd Edition, Cambridge university press, Cambridge 1991, qui nécessitent un isolement préalable du micro-organisme en culture pure, puis l'étude de ses caractéristiques morphologiques et surtout physiologiques (plus de 80 tests), l'ensemble de ces opérations requérant de 10 à 30 jours pour une identification. Ces méthodes sont donc fastidieuses et particulièrement longues. Elles sont réservées aux laboratoires de référence et inutilisables en pratique courante.

Plus rapide et plus rationnelles, les méthodes miniaturisées et standardisées d'identification des levures d'importance médicale, comprennent cependant toujours un isolement préalable puis une identification du micro-organisme à l'aide de tests physiologiques, réalisés par exemple à l'aide de galeries. La durée de tels tests est d'environ 5 jours pour aboutir à une identification. De plus ces techniques donnent parfois des résultats équivoques. Des espèces telles que *C. krusei*, *C. inconspicua*, *C. lipolytica*, *C. norvegensis*, *C. rugosa* et *C. valida* sont parfois, voire même souvent, confondues.

Avec le progrès de la biologie moléculaire, une nouvelle technologie utilisant le principe d'hybridation ADN-ADN ou ADN-ARN s'est développée pour la mise au point des tests d'identification rapides, sensibles et spécifiques. Le matériel génétique des micro-organismes présents dans l'échantillon est décelé directement à l'aide de sondes nucléiques à ADN ou ARN marqué. La détection et l'identification peuvent être réalisées simultanément sans isolement préalable.

Les rares méthodes décrites d'identification des levures à l'aide de sondes nucléiques, présentent l'inconvénient d'être basées sur des sondes de petites tailles (30 à 35 nucléotides) qui nécessitent la mise en oeuvre d'un marquage radioactif plus sensible que le marquage non radioactif mais aussi plus contraignant.

5 Un autre inconvénient de ces sondes lié à leur taille réduite, s'exprime dans le cadre de leur mise en oeuvre dans les techniques de polymérisation en chaîne (PCR). Ces dernières consistent à amplifier une séquence ADN donnée, en la multipliant à partir d'un couple d'amorces et l'aide d'une polymérase. De telles amplifications peuvent être employées dans un but diagnostique, afin de faciliter la détection. Dès  
10 lors que ces sondes seraient utilisées comme amorces en PCR cela limiterait les possibilités de choix des amorces dans une petite séquence de 30 à 35 nucléotides.

Ces sondes sont généralement dirigées contre des séquences cibles d'ADN codant la petite sous unité 18 S de l'ARN ribosomique (ssuARNr). Ces cibles sont classiquement choisies car les séquences d'ADN codant les ARN ribosomaux sont les  
15 premières à avoir été déterminées. De même l'utilisation de la partie 3' des séquences codant la grande sous unité 25 S de l'ARN ribosomique (lsuARNr) a été décrite. Il existe un problème lié au choix de séquences codant les ssuARNr ou lsuARNr comme cibles. En effet ces séquences ont un caractère universel, c'est-à-dire qu'une partie importante d'entre elles est retrouvée chez tous les micro-organismes (levures,  
20 bactéries) ; elles sont dites "hyper-conservées". Le choix des séquences cibles est donc limité aux portions limitées variables de ces séquences codantes des ARNr.

En outre, parmi les sondes décrites dans l'art antérieur, aucune ne permet le typage des souches, c'est-à-dire la différenciation des individus au sein de l'espèce, ces individus étant uniquement assignés à un groupe : e.g. l'espèce *Candida krusei*,  
25 l'espèce *Candida glabrata*, l'espèce *Candida lusitanae*, etc... En effet, les sondes préalablement décrites ne permettent pas un typage des souches au sein des espèces car elles ne révèlent pas suffisamment de polymorphisme de la taille des fragments de restriction parmi les souches de ces espèces et ainsi ne constituent pas de marqueurs épidémiologiques fiables.

30 A titre d'illustrations de telles sondes connues et répondant imparfaitement aux besoins techniques existant en la matière, on peut citer :

- les sondes oligonucléotidiques développées par la société Genetrak Systems et dirigées contre des séquences de 30 (probes 1351), 33 (probe 1537) et 35 (probe 1530) nucléotides de l'ADN codant l'ARN ribosomal 18 S. Ces sondes sont  
35 décrites dans la demande de brevet européen N° 0 422 869. Elles ont été testées contre 4 souches de *Candida krusei*. Il est à noter qu'il n'est pas fait état dans les

exemples de cette demande, de détection de *Candida krusei* par hybridation dans les prélèvements de sang humain, crachats ou liquide céphalorachidien. Par ailleurs, ces sondes ne permettent pas de différencier les souches entre elles au sein de l'espèce *Candida krusei*.

- 5           •       la sonde décrite par Niesters et collaborateurs (Niesters et al., 1993, Rapid polymerase chain reaction-based identification assays for *Candida* species. J. Clin. Microbiol., 31, 904-910) 1993). Il s'agit d'une sonde de 20 nucléotides (oligonucléotide 705) dirigée contre une séquence spécifique du produit d'amplification de la petite sous unité d'ARN ribosomal (ssuARNr). Cette étude est  
10 utilisée dans une méthode d'identification des espèces du genre *Candida* basée sur l'amplification en chaîne (PCR) de séquences du gène codant la ssuARNr, puis le séquençage direct de ces séquences. Cette méthodologie concerne, là encore, l'ARN ribosomal. Elle présente l'inconvénient de nécessiter un séquençage direct du produit d'amplification, technique qui n'est accessible qu'à un nombre limité de laboratoires de  
15 recherche et surement pas applicable en routine. De plus, d'après les auteurs, l'oligonucléotide 705 ne permet pas une amplification spécifique de l'ADN de *Candida krusei*. Les sondes décrites par Niesters et collaborateurs ne permettent pas elles non plus de typer les souches de *Candida krusei*.

- On connaît également, notamment au travers de l'article rédigé par BALEIRAS  
20 COUTO et al - Applied & Environmental Microbiology, Jan. 1996 p.41-46 "Evaluation of molecular typing techniques to assign genetic diversity among *Saccharomyces cerevisiae* strains". Cet article divulgue la discrimination de souches à l'intérieur de l'espèce *Saccharomyces cerevisiae*, par recours aux techniques d'identification du type de celles faisant intervenir des fragments d'ADN  
25 polymorphiques amplifiés de manière aléatoire (RAPD) et du type réaction de polymérisation en chaîne (PCR) à l'aide d'amorces nucléotidiques. Les fragments obtenus par ces techniques sont ensuite soumis à des digestions à l'aide d'enzyme de restriction. Toutes les séquences et fragments ainsi obtenus sont testés dans l'identification infraspécifique de 16 souches différentes de *Saccharomyces cerevisiae*.  
30 La RAPD et la PCR ont permis de reconstituer les régions d'espacements intergéniques transcrites (IGR - ITS) et les régions intergéniques non transcrites (IGR - NTS) de cistrons d'ADNr codant pour la synthèse d'ARN ribosomique (ARNr). Ce document ne décrit nullement le contenu des séquences reconstituées par RAPD et PCR dont la longueur est d'ailleurs d'au moins 7 kbases (kb). Cet article enseigne qu'il  
35 n'existe pas de polymorphisme significatif entre les souches de *Saccharomyces cerevisiae* A. Dans le meilleur des cas (correspondant à celui où l'on utilise des *TaqI*),

il n'est possible de discriminer que 6 souches sur 16. Ce résultat est tout-à-fait perfectible.

Par ailleurs, BALEIRAS COUTO et al incitent plutôt le lecteur à s'intéresser aux régions intergéniques d'espacement transcrite (ITS), pour la différenciation infraspécifique de levure. Il présente ces régions intergéniques transcrites comme

étant plus appropriées que les régions intergéniques non transcrites (NTS). En outre, les amorces (Primers) utilisées dans la technique de (PCR) préconisées dans ce document, ne fonctionnent pas de façon universelle sur toutes les espèces de levures, ceci constitue un facteur limitatif de la technique. En particulier les auteurs

indiquent que l'amplification n'a pas été possible pour les espèces *C. valida*, *C. krusei*, *Z. bailii*, ou *Z. rouxii*.

Il apparaît donc que les techniques décrites dans cet article ne conviennent que pour du typage infraspécifique de levure et notamment pas pour l'identification d'espèces de levure.

La demande de brevet internationale WO 95/11 991 décrit une sonde et un procédé de détection des levures de l'espèce *C. krusei*. Les outils génétiques constituant cette sonde sont formés par un fragment d'ADN et/ou d'ARN de taille comprise entre 7 et 4 kb, hybridant spécifiquement avec l'ADN et/ou l'ARN de *Candida krusei* et ne codant pas pour l'ARN ribosomal (ARNr).

Sont également compris dans ces outils génétiques, les produits de transcription et de traduction du susdit fragment, de même que les associations entre ces différents outils. Cette demande de brevet ne divulgue pas la structure des fragments d'ADN mis en oeuvre dans cette sonde. Le moyen de caractérisation utilisé pour ces fragments, est formé par leur carte de restriction enzymatique. L'espèce de levure exclusivement

concernée par ce document est *Candida krusei*.

La demande PCT WO-A-93 23 568 (A.R. HOLMES et al) concerne des méthodes de diagnostic des infections fongiques, selon lesquelles on met en oeuvre des sondes nucléiques issues de fragment d'ADN provenant exclusivement de *C. albicans* et exclusivement applicables à cette espèce de levure. Dans cette demande de brevet, on préconise de manière préférée la mise en oeuvre d'une sonde EOB2 constituée par tout ou partie d'une séquence nucléotidique, comprise dans le gène d'ADNr codant pour la sous-unité d'ARNr 5S. Cette séquence d'ADNr comprend 121 paires de bases localisées entre le nucléotide 937 et le nucléotide 1057 de la SEQ ID N° 1 du gène d'ADNr 5S, divulguée dans cette demande PCT.

Secondairement, le WO-A-93 23 568 décrit une sonde EOB1 localisée dans la région intergénique non transcrite (IGR-NTS), située entre le gène d'ADNr codant pour

l'ARNr 5S et le gène d'ADNr codant pour l'ARNr 18 S. Cette sonde EOB1 s'hybride avec une séquence d'ADNr comprise entre les nucléotides 1 et 936, de préférence 1 et 710 de la SEQ ID NO : 1.

5 Dans le cas où elles sont de faible longueur, ces sondes EOB2 et EOB1 peuvent être amplifiées par PCR amorces PS pA<sub>1</sub>, PS pA<sub>2</sub>, PCon<sub>1</sub>, PCon<sub>2</sub>.

La sonde EOB2 comprise dans l'ADNr 5S permet la détection des levures pathogènes, tandis que la sonde EOB1 est présentée comme étant interspécifique. En tout état de cause, il est clair que le WO-9323568 ne fait nullement mention de distinction infra-spécifique des souches des différentes espèces de levures citées  
10 (*C. albicans*, *C. krusei*, etc). La sonde EOB1 n'est pas une séquence source de structure particulière, comme e.g. de type itérative comprenant plusieurs sous-séquences répétées homologues entre elles.

Le WO-93 23 568 ne divulgue donc pas, de structure génique typique d'identification et de détection inter et infraspécifique.

15

#### BREVE DESCRIPTION DE L'INVENTION :

Dans cet état de la technique, l'un des objectifs essentiels de la présente invention est de fournir une sonde de détection et d'identification spécifique et/ou  
20 infraspécifique de levure, à vocation universelle dans le domaine des levures.

Un autre objectif essentiel de l'invention est de fournir une sonde, notamment nucléique, qui permette à la fois une détection spécifique et une identification infra spécifique de levure, de manière performante, fiable et reproductible.

Un autre objectif de l'invention est de fournir une sonde, notamment nucléique,  
25 qui soit de taille suffisante pour autoriser l'utilisation d'un marqueur, autre que radioactif et donc de mise en oeuvre plus commode.

Un autre objectif de l'invention est que cette sonde nucléique soit dirigée contre des séquences cibles d'ADN différent de celui codant pour des petites sous-unités 5,8 S ou 18 S d'ARN ribosomique, qui sont globalement peu représentatifs de l'espèce,  
30 car très conservées chez les micro-organismes.

Un autre objectif de l'invention est que cette sonde puisse être utilisée dans des milieux complexes biologiques du type sang humain, crachat, ou liquide céphalorachidien.

Un autre objectif essentiel de l'invention est de fournir une sonde notamment à  
35 *Candida krusei* qui permette de dépasser la simple détection d'espèces pour accéder à



l'identification précise et au typage des levures au sein même de l'espèce *Candida krusei*.

Un autre objectif essentiel de l'invention est de fournir une sonde notamment à *Geotrichum candidum* qui permette de dépasser la simple détection d'espèces pour  
5 accéder à l'identification précise et au typage des levures au sein même de l'espèce *Geotrichum candidum*.

Un autre objectif essentiel de l'invention est de fournir une sonde de détection de levure qui soit composée d'outils génétiques du type séquence d'ADN et/ou d'ARN parfaitement identifiée et reproductible.

10 Un autre objectif essentiel de l'invention est de fournir une sonde de détection de levure qui soit significativement performante pour l'identification inter et infraspécifique d'un suffisamment grand nombre de levures, pour pouvoir être économique viable.

Un autre objectif essentiel de la présente invention est de fournir un procédé de  
15 détection et d'identification inter/infraspécifique de levures qui soit simple, non délicat et économique à mettre en oeuvre et dans lequel on a recours à la sonde sus-visée.

Un autre objectif essentiel de l'invention est de fournir une sonde applicable notamment à la détection spécifique à l'identification infraspécifique de levure mais également comme marqueur épidémiologique, comme traceur de souches, ou dans le  
20 cadre d'une stratégie thérapeutique visant les affections à levure, entre autres.

Ces objectifs, et d'autres encore, sont atteints par la présente invention qui concerne une sonde, notamment, de détection spécifique et/ou intraspécifique de levures, du type de celle choisie parmi les outils génétiques (ou apparentés) suivants :

- au moins une partie :

- 25                   \*       d'au moins un fragment F d'acide nucléique ADN et/ou ARN recopié, construit et/ou isolé à partir d'au moins une séquence-source d'ADN ribosomal (ADNr) :
- ne codant pas pour la synthèse d'ARN ribosomal (ARNr),
- 30                   →       et localisée dans la région d'espacement intergénique (IGR) non transcrite (NTS) et comprise entre l'ADNr codant pour la sous-unité 25 S d'ARNr, et l'ADNr codant pour la sous-unité 18 S d'ARNr,
- 35                   \*       et/ou d'au moins un analogue Fa de ce fragment, résultant de la dégénérescence du code génétique ;

- \* et/ou d'au moins un fragment Fc d'ADNc complémentaire du fragment F,
  - au moins une partie des produits de transcription primaire du fragment F et/ou Fa et/ou Fc.
  - 5 - et une association des outils énumérés ci-dessus.
- caractérisé en ce que la séquence-source d'ADNr pour F et/ou Fa et/ou Fc est sélectionnée de telle sorte
- qu'elle soit itérative
  - qu'elle comprenne au moins 2, de préférence de 2 à 12 sous-séquences répétées,
  - 10 ➤ et que le degré d'homologie entre ces sous-séquences soit supérieur ou égal à 50 %, de préférence 55 %, et plus préférentiellement encore à 60 %.

15

#### DESCRIPTION DETAILLÉE DE L'INVENTION :

Il est donc du mérite des inventeurs cités dans la présente demande d'avoir pu isoler et caractériser un certain nombre de fragments F d'acide nucléique ayant une structure tout-à-fait particulière et qui *de facto*, constitue une véritable empreinte digitale non seulement d'une espèce donnée de levure mais également des différentes souches contenues dans cette espèce. Le mérite des inventeurs est d'autant plus grand que l'isolement et la caractérisation desdits fragments a été réalisé au sein des zones d'espacement intergénique non transcrites de cistrons ribosomaux, alors qu'il existait un courant scientifique, illustré notamment par l'article de BALEIRAS COUTO et al visé ci-dessus, selon lequel il semblait préférable et recommandé de s'intéresser aux régions transcrites des gènes ribosomaux pour la détection et l'identification de levures.

Il relève également du mérite des inventeurs d'avoir proposé des outils génétiques ou apparentés constitués par les susdits fragments pour la détection et l'identification inter-espèce de levures, de même que pour la mise en évidence infraspécifique (discrimination des levures d'une espèce donnée) et, en particulier, sans que cela ne soit limitatif, des espèces *Candida krusei* et *Geotrichum candidum*

Les fragments F, Fa, Fc d'acide nucléique selon l'invention ont pour caractéristique originale et avantageuse d'être non transcrits et de comprendre un ou plusieurs modules séquentiels de bases nucléiques (dénommés également séquences-sources itératives dans le présent exposé) comportant chacun une ou plusieurs sous-

séquences répétées qui sont des copies très proches les unes des autres. C'est ou ce sont ces modules qui constituent un ou plusieurs motifs primaires utiles pour la détection/identification inter et infraspécifique d'une pluralité de levures.

Le nombre de bases nucléotidiques de ces modules ou séquences-sources varient dans une espèce à l'autre, mais on retrouve cette organisation similaire récurrente dans de multiples espèces de levures.

Sans vouloir être lié par la théorie, il semble que le moyen d'identification infraspécifique des sondes conformes à l'invention, puisse être fondé sur le polymorphisme qui, au sein d'une espèce donnée de levure, découle du nombre de sous-séquences dans les modules ou séquences-sources itératives, ou entrant dans la constitution des fragments F, Fa, Fc des sondes selon l'invention.

Suivant un premier mode de réalisation de l'invention, premier mode qui est d'ailleurs préféré, la séquence-source de F est localisée dans la région d'espacement intergénique 2 (IGR<sub>2</sub>) non transcrite (NTS) et comprise entre l'ADNr codant pour la sous-unité 5S d'ARNr et l'ADNr codant pour la sous-unité 18 S d'ARNr (Cf Fig.2 annexée).

Selon une variante avantageuse du premier mode de réalisation de l'invention, la séquence-source itérative de F correspond à une séquence F $\alpha$  et comprend au moins trois sous-séquences répétées comportant chacune au moins 80, de préférence au moins 120 et plus préférentiellement encore entre 150 et 300 bases nucléotidiques. En pratique, la séquence-source F $\alpha$  comprend entre 3 et 10 sous-séquences répétées comprenant entre 150 et 250 bases nucléotidiques.

De préférence, les séquences-sources itératives de F et, en particulier F $\alpha$ , de la sonde selon l'invention comprennent chacune au moins une sous-séquence de longueur tronquée par rapport aux autres sous-séquences, cette sous-séquence tronquée étant de préférence terminale.

S'agissant de la séquence-source itérative F $\alpha$ , il est à considérer, sans pour autant vouloir être lié par la théorie, que le nombre de sous-séquences répétées élémentaires détermine le polymorphisme au sein d'une même espèce de levures. Cela pourrait résulter d'une réplication (duplication d'une ou plusieurs des sous-séquences répétées). Le tronc commun structurel entre les différentes souches d'une même espèce tient à une similitude dans l'organisation des répétitions de chaque sous-séquence.

Les phénomènes de réplication (duplication) à l'origine du polymorphisme entre souches sont liés à des recombinaisons, dont l'éventuelle séquence tronquée terminale évoquée ci-dessus, serait le reflet.

Toujours dans le cadre du mode préféré de mise en oeuvre de l'invention, selon lequel la séquence source F est dans l'IGR<sub>2</sub>, il est prévu, conformément à une autre variante avantageuse, que la séquence-source itérative de F correspond à une séquence F $\beta$  et comprend au moins trois sous-séquences répétées comportant

5 chacune au moins 4, de préférence au moins 6 et, plus préférentiellement encore entre 7 et 25 bases nucléotidiques, chaque sous-séquence comprenant au moins un motif commun (Com 1) à toutes les sous-séquences et ayant une longueur inférieure ou égale à celle de la sous-séquence à laquelle il appartient.

En pratique, ces séquences-sources F $\beta$  comprennent chacune 8 à 12 sous-séquences

10 répétées comportant individuellement entre 7 et 13 bases nucléotidiques.

Conformément au deuxième mode de réalisation de la sonde selon l'invention, la séquence source de F :

- est localisée dans la région d'espace intergénique 1 (IGR<sub>1</sub>) non transcrite (NTS) et comprise entre l'ADNr codant pour la sous-unité 5S d'ARNr et l'ADNr

15 codant pour la sous-unité 25 S d'ARNr

- et correspond à au moins une séquence F $\gamma$  et comprend au moins trois sous-séquences répétées comportant chacune au moins 4, de préférence au moins 5 et, plus préférentiellement encore entre 5 et 50, voire entre 5 et 45 et mieux encore entre 5 et

20 40 bases nucléotidiques, chaque sous-séquence comprenant au moins un motif commun (Com 2 ou 3) à toutes les sous-séquences et ayant une longueur inférieure ou égale à celle de la sous-séquence à laquelle il appartient.

un cas de figure particulièrement préféré étant celui dans lequel F correspond à au moins deux séquences F $\gamma$  : F $\gamma_1$  et F $\gamma_2$  ayant respectivement avantageusement de 5 à 20 et de 9 à 25 bases nucléotidiques.

25 Selon un autre de ces aspects, la présente invention concerne également des séquences d'ADN, de préférence ribosomique (ADNr), susceptibles de correspondre aux fragments F sus-évoqués : F $\alpha$ , F $\beta$ , F $\gamma$ , F $\gamma_1$  et F $\gamma_2$ , ou susceptibles de contenir tout ou partie d'un ou plusieurs desdits fragments F noyés au sein d'une longue séquence nucléique, de préférence d'ADNr.

30 La présente invention a donc ainsi également pour objet une séquence d'ADNr caractérisée en ce qu'elle est choisie dans le groupe des séquences annexées suivantes : SEQ ID NO : 1 ; SEQ ID NO : 2 ; SEQ ID NO : 3 ; SEQ ID NO : 4 ; SEQ ID NO : 5 ; les SEQ ID NO : 2 : à 5 : étant particulièrement préférés.

En l'occurrence, la SEQ ID NO : 1 annexée contient toutes les SEQ ID NO 2

35 à 5 :

- SEQ ID NO : 2 :  $\rightarrow$  base N 2256 à base 3512

- SEQ ID NO : 3 : → base N 4823 à base 4903
- SEQ ID NO : 4 : → base N 923 à base 1011
- SEQ ID NO : 5 : → base N 1012 à base 1151

Selon un exemple particulier d'illustration de l'invention, la sonde qu'elle  
5 concerne est caractérisé en ce que :

F $\alpha$  correspond à SEQ ID NO : 2 :

F $\beta$  correspond à SEQ ID NO : 3 :

F $\gamma_1$  correspond à SEQ ID NO : 4 :

F $\gamma_2$  correspond à SEQ ID NO : 5 :

10 Cet exemple de réalisation correspond à une souche de *Candida krusei*.

Suivant une disposition ou une modalité intéressante de l'invention, la sonde décrite ci-dessus est caractérisée

- en ce qu'elle comprend au moins une partie d'au moins l'un des fragments F $\alpha$ , F $\beta$ , F $\gamma_1$ , F $\gamma_2$ , le fragment F $\alpha$  étant particulièrement préféré.

15 - en ce qu'elle est appliquée :

- ▷ à la détection/identification spécifique et infraspécifique de levures,
- ▷ ou comme marqueur épidémiologique,
- ▷ ou encore comme traceur de souche,
- ▷ ou bien enfin dans le cadre d'une stratégie thérapeutique visant des

20 affections à levures.

S'agissant de l'origine des séquences d'acides nucléiques typiques, conformes à l'invention, on a vu qu'elles sont naturelles puisqu'issues à la base du génôme de souches de levures.

25 Mais il va de soi qu'il est tout à fait à la portée de l'homme du métier, de préparer des séquences nucléiques « non naturelles » en recopiant les séquences par exemple selon les principes du clonage ou de la réaction de polymérisation en chaîne (PCR), ou par voie de synthèse. Notamment, on peut synthétiser des séquences nucléiques à l'aide d'appareils de synthèse automatiques, tels que ceux commercialisés par Applied Biosystems.

30 Aussi, dans tout le présent exposé les termes fragments ou séquences nucléiques, désignent aussi bien ceux de type synthétique, que ceux de type naturel.

Dans le cadre de l'application en détection / identification, les fragments F, Fa et Fc sont utilisés comme sonde selon le principe d'hybridation ADN-ADN ou ADN-ARN. Une sonde constituée de tels fragments permet la réalisation de tests de  
35 détection et d'identification rapides, sensibles, et inter/infraspécifiques d'une pluralité d'espèces de levure. En d'autres termes, cette sonde peut notamment constituer un

excellent marqueur de typage des souches d'une espèce donnée et en particulier de *Candida krusei* ou *Geotrichum candidum*, ces dernières espèces ayant été sélectionnées, parmi d'autres, pour servir de support à l'exemplification du présent exposé.

5 Il faut considérer que l'hybridation de la sonde nucléique de l'invention avec des séquences cibles d'acide nucléique de levure (e.g. *C. krusei*), peut se faire à l'aide de la totalité des fragments F visés ci-dessus mais aussi avec seulement une fraction de tout ou partie de ces fragments, dont la taille reste, de préférence, supérieure ou égale à 5 Bases, de préférence à 10 Bases et, plus préférentiellement encore à 100 Bases.

10 Une autre des caractéristiques avantageuses de la sonde nucléique selon l'invention est liée au fait que cette dernière peut être marquée à l'aide de moyens de marquage aptes à révéler l'hybridation. Ces moyens de marquage peuvent être radioactifs ou non. Parmi les marqueurs radioactifs classiques, on peut citer le  $P^{32}$ . S'agissant des marqueurs non radioactifs, on peut mentionner à titre d'exemples, les  
15 enzymes fixés tels que la peroxydase.

Par ailleurs, la sonde selon l'invention peut éventuellement être employée comme source d'amorces spécifiques utilisables dans des réactions de polymérisation en chaîne (PCR). Ces amorces constituent un autre objet de l'invention.

Comme cela a déjà été signalé, la sensibilité de la sonde selon l'invention est  
20 excellente, de sorte qu'elle est parfaitement adaptée à la détection et/ou au typage de levures (e.g. *Candida krusei* ou *Geotrichum candidum*) par hybridation en "dot-blot" ou en "southern-blot". Cette sensibilité se vérifie dans des essais indirects d'hybridation (méthode des dépôts, dot-blot, Southern-blot, Northern-blot, technique d'hybridation « sandwich ») mais aussi dans des expériences d'hybridation directe in  
25 situ. Ces dernières consistent à faire une empreinte sur filtre, de colonies en croissance sur un milieu de culture solide et à amener la sonde d'hybridation directement sur le filtre.

Comme cela ressort du texte supra ainsi que de la revendication 1, la sonde selon l'invention est de préférence de nature nucléique, c'est-à-dire constituée d'acides  
30 nucléiques du type ADN, ADNc, ARNm ou leurs variants et analogues. Ces acides peuvent être d'origine naturelle ou synthétique.

Il est intéressant de noter que la sonde selon l'invention peut être employée efficacement dans tout milieu complexe, comme peuvent l'être les milieux biologiques : sang, crachat, liquide céphalorachidien, etc. Cela n'interfère nullement  
35 sur ses qualités de détection et d'identification.

La présente invention concerne également un procédé de détection et d'identification inter et/ou infraspécifique de levure caractérisé en ce qu'il consiste à mettre en oeuvre au moins une sonde nucléique telle que décrite ci-dessus.

Ce procédé s'inscrit dans le cadre des méthodologies connues dans le domaine de la détection et de l'identification génétique des microorganismes.

De préférence, ce procédé comprend les étapes suivantes :

- on extrait l'ADN total génomique des souches à étudier,
- on soumet éventuellement cet ADN total à une digestion enzymatique à l'aide d'au moins une enzyme de restriction,
- on dénature l'ADN total éventuellement digéré,
- on met en présence l'ADN total ainsi dénaturé, avec la sonde dotée d'au moins un marqueur, de façon à réaliser l'hybridation,
- on élimine l'ADN et la sonde non hybridée,
- et on révèle l'hybridation à l'aide du marqueur.

Il s'agit d'une technique d'hybridation dans laquelle l'ADN cible du micro-organisme à étudier, est soumis à une dénaturation, avec ou sans digestion enzymatique préalable. L'étape qui suit est celle de réassociation des brins séparés d'ADN dénaturé avec la sonde pour reformer des appariements de paires de base originaux.

Lors de cette étape de réassociation, les molécules simple brin de la sonde et de la cible sont mises en condition plus ou moins favorables pour l'appariement (plus ou moins stringentes). Dans des conditions très stringentes seules les molécules dont les séquences sont complémentaires pour un grand nombre de bases, s'hybrident pour former une molécule double brin. La sonde est alors spécifique de la cible.

Avec une sonde marquée à l'aide d'un élément radioactif, tel que le  $P^{32}$ , ou d'une enzyme greffée comme par exemple la peroxydase, l'hybridation est aisément révélée qualitativement et quantitativement.

Dans le cas où la sonde comprend des fragments (séquences) nucléiques typiques d'origine naturelle, ceux-ci sont naturellement dénaturés à l'instar des cibles qu'ils sont susceptibles de détecter. Ce n'est évidemment pas le cas pour des fragments synthétiques.

La présente invention vise également un réactif de détection/identification de levures comprenant au moins une sonde selon l'invention.

## APPLICATION INDUSTRIELLE

Il ressort de ce qui précède que la sonde nucléique selon l'invention ainsi que le procédé en faisant application, sont parfaitement spécifiques des levures (e.g. *Candida krusei* ou *Geotrichum candidum*) et offrent une excellente sensibilité. Cela leur ouvre des débouchés tout à fait intéressants dans les domaines applicatifs de l'identification et du criblage différentiation des souches de différentes espèces. (diagnostic médical, contrôles industriels en alimentaire, fermentation...).

La sonde et le procédé selon l'invention intéressent diverses espèces de levure et, en particulier, celles d'intérêt médical telles que celles du genre *Candida* ou d'intérêt industriel (agroalimentaire tel que *Geotrichum Candidum*).

Outre, la détection qui regroupe l'identification et le typage de microorganisme de type levure, la sonde selon l'invention pourrait être appliquée dans le cadre d'une stratégie thérapeutique dirigée à l'encontre des affections à levures. Plus précisément, cela signifie que les fragments F formant la sonde peuvent être employés comme cibles à atteindre, pour des principes médicamenteux actifs contre les levures. Ce rôle de cible pour médicament pourrait notamment être tenu par l'ADN ou par ses produits de transcription primaire.

L'emploi de la sonde selon l'invention comme marqueur épidémiologique ou comme traceur de souches constituent d'autres variantes d'applications de la sonde selon l'invention.

L'illustration de la présente invention effectuée ci-après sur la base de travaux réalisés pour les espèces de levures *Candida krusei* et *Geotrichum candidum*, mais il va de soi que les exemples basés sur *Candida krusei* et *Geotrichum candidum* pourront être extrapolés aux autres espèces de levures et, notamment à celles d'intérêt industriel et médical.

En tout état de cause, l'invention sera mieux comprise, d'autres de ses avantages et variantes de réalisation ressortiront bien, à la lumière des exemples non limitatifs qui suivent et qui décrivent en référence aux dessins annexés la structure d'une sonde selon l'invention et en particulier à des fragments F qui la constitue, ainsi que plusieurs procédures de détection et d'identification en faisant application.

## DESCRIPTION DES FIGURES :

- la liste des séquences annexée donne la structure des SEQ ID : 1 à 5 : correspondant respectivement aux fragments F1 et F $\alpha$ , F $\beta$ , F $\gamma_1$  et F $\gamma_2$ .



Le fragment F1 donne le séquençage de la région intergénique d'espacement (IGR) entre l'ADNr 25S et l'ADNr 18 S d'un cistron ribosomal de *Candida krusei*.

Les exemples de fragment F $\alpha$ , F $\beta$ , F $\gamma_1$  et F $\gamma_2$  seront également dénommés respectivement CKRS1/ b / a1 / a2, dans ce qui suit.

- 5 - la Figure 1 annexée représente l'hybridation Southern-blot de *Candida krusei* LMCK31 par des sondes F1 et/ou F2.
- la Figure 2 est une représentation schématique d'un cistron d'ADN ribosomal (ADNr) montrant la localisation du fragment F1 dans la région d'espacement intergénique (IGR) et, au sein de cette dernière, l'existence des fragments F $\alpha$ , F $\beta$ , (IGR $_2$ ) de même  
10 que F $\gamma_1$  et F $\gamma_2$  (IGR $_1$ ), conformes à l'invention.
- la Figure 3 est une carte de restriction du fragment F1 de la Figure 2.
- la Figure 4 est une représentation schématique de la stratégie de séquençage, en corrélation avec le fragment F1 de la figure 3.
- la Figure 5 est une représentation schématique éclatée du fragment F $\alpha$  montré aux  
15 Figures 2 et 3.
- la Figure 6 donne le séquençage de F $\alpha$ , en particulier de ses sous-séquences Kre-0 à 7, en faisant une comparaison entre les différentes sous-séquences. Des étoiles représentent des bases analogues à celles de la sous-séquence Kre-1. >>ivrl<< est une séquence inversée comprenant 8 bases. <<pal1>> est un Palindrome de 8 bases.
- 20 - Les Figures 7A, 7B et 7C sont des représentations comparées des sous-séquences formant les séquences-sources itératives F $\beta$ , F $\gamma_1$  et F $\gamma_2$  respectivement.
- les Figures 8A et 8B sont des électrophorèses à champ pulsé de chromosomes de *C. krusei* et de *C. albicans*, respectivement avant et après hybridation avec des sondes  
F1,  
25 F $\alpha$  et F $\beta$ , F $\gamma_1$  et F $\gamma_2$ .
- la Figure 9A est une photographie d'une séparation par électrophorèse sur gel d'agarose, de fragments d'ADN de souches de différentes espèces de levures (1 à 11), digérés par *EcoRI*.
- la Figure 9B est une hybridation selon Southern d'ADN de différentes espèces de la  
30 Figure 9A, illustrant la spécificité de la séquence-source itérative F $\alpha$  (ou CKRS-1) utilisée comme sonde.
- la Figure 10A est une photographie du gel d'électrophorèse de l'ADN de plusieurs souches de l'espèce *Candida krusei*.
- la Figure 10B est une hybridation selon Southern de l'ADN de différentes souches de  
35 *C. krusei*, illustrant le polymorphisme mis en évidence lorsque la séquence F $\alpha$  (ou CKRS-1) est utilisée comme sonde.

- la Figure 11 montre le résultat de l'amplification sélective en PCR de la région comprenant F $\alpha$ .
- la Figure 12 est une représentation schématique de la Carte de restriction du Fragment F<sub>10</sub> forment la sonde A<sub>2</sub>O<sub>27</sub> de *Geotrichum candidum*. Cette sonde  
5 comprenant un sous-fragment F dont la séquence-source itérative est F $\delta$  (région variable GCRS-A).
- la Figure 13 représente un profil d'hybridation (Southern) avec la sonde A<sub>2</sub>O<sub>27</sub> spécifique de *Geotrichum candidum* et développée selon le mode de la sonde F(F $\alpha$ , F $\beta$ , F $\gamma_1$ , F $\gamma_2$ ) de *Candida krusei*.

## 10 EXEMPLES

EXEMPLE I : SEQUENÇAGE D'UN FRAGMENT F1 D'ADNr DE *C. KRUSEI* PROVENANT DE LA DIGESTION DE LA SOUCHE LMCK1 PAR *ECORI*

- 15 1.1. La souche LMCK 31 a été isolée dans un prélèvement broncho-alvéolaire chez un patient hospitalisé. Cette souche a été déposée le 22 Octobre 1993 à la CNCM de l'Institut Pasteur de PARIS sous la référence I-1372.
- Cette souche LMCK 31 est cultivée sur des milieux de culture du type gel AGAR YM (1% glucose, 0,3 % extrait levure, 0,3 % extrait de malt, 0,5 % de pactopeptone) ou  
20 AGAR SM (1% glucose, 1% bactopeptone, 0,1 % MGSO<sub>4</sub>. 7H<sub>2</sub>O, 0,22 % KH<sub>2</sub>PO<sub>4</sub>, 0,1 % K<sub>2</sub> HPO<sub>4</sub>, 0,1 % extrait levure).

### 1.2. Isolement et clonage d'un fragment F1

- Un fragment F1 de longueur approximativement égale à 5,4 kb est isolé à partir  
25 d'ADN génomique de la souche LMCK31, conformément au protocole décrit dans la demande de brevet PCT WO-95 11 991. Le fragment F1 a été sélectionné parmi la multitude de fragments produits par la digestion de l'ADN total de LMCK31 par *EcoRI* ou par *HinfI*.
- Plus précisément on a réalisé une hybridation "southern blot" en mettant en oeuvre 5  
30  $\mu$ g d'ADN génomique de *Candida krusei* LMCK31 digéré soit par *EcoRI* ou *HinfI* puis séparé par électrophorèse selon leur taille sur un gel d'agarose (0,8 % m/v). L'ADN séparé est dénaturé dans le gel par une solution de soude puis neutralisé par un tampon tris. L'ADN est ensuite transféré sur une membrane en nylon (Amersham life science inc., ARLINGTON HEIGHTS, IL) puis hybridé avec la sonde contenant des  
35 fragments F1 et F2 correspondant à l'IGR1 et 2 d'ADNr ou avec des plasmides

contenant, soit le fragment F1, soit le fragment F2. La révélation de l'hybridation par "southern-blot" est effectuée à l'aide du kit ECL d'Amersham. Les bandes d'hybridation sont obtenues par marquage avec la peroxydase et révélation par chimiluminescence.

- 5 La Figure 1 annexée donne les résultats obtenus. La colonne A correspond aux produits de la digestion par *EcoRI* de l'ADN de *Candida krusei* non hybridé avec une sonde. La colonne B correspond à l'hybridation avec une sonde comprenant des fragments F1 et F2. La colonne C correspond à une hybridation avec une sonde plasmidique contenant F1 et la colonne D à l'hybridation avec une sonde plasmidique
- 10 contenant F2.

Les poids moléculaires estimés des quatre bandes d'hybridation des colonnes B, C, D sont données en kilobase (kb) en marge gauche.

### 1.3. Séquençage de F1

- 15 F1 est séquencé dans son intégralité dans les deux directions. La méthode utilisée pour ce séquençage est une méthode classique dans le domaine technique considéré. Il s'agit de la méthode par les terminaisons DIDEOXY utilisant la séquenase version 2.0. (US biochemical, cleveland OH). Les moyens utilisés sont le kit ERASE-A-base (Promega - Madison, WI).

- 20 La Figure 2 montre de manière schématique la localisation du fragment F1 dans le cistron d'ADN ribosomal de LMCK 31.

La Figure 3 annexée donne la carte de restriction du fragment F1. Les abréviations utilisées pour les enzymes de restriction ont les significations suivantes : E = *EcoRI* (0,5261), Ev = *EcoRV* (1803), H = *Hinfi* (229, 1155; 1551, 4828, 4864), N = *NsiI* (401, 4383), P = *PstI* (2183), S = *SmaI* (803); St = *StyI* (2425, 2589, 2754, 2919, 3084, 3249, 3414).

- La Figure 4 annexée donne la stratégie de séquençage adoptée. Dans cette Figure 4, les chiffres encadrés 1, 2, 3 correspondent à des sous-clones produits de délétions. Les chiffres donnés entre parenthèses après les enzymes correspondent au numéro de
- 30 base nucléotidique identifiant le site de coupure.

Il apparaît donc que F1 est une macro-séquence comprenant 5261 paires de base et ayant un taux en bases GC de 42,2 % molaire. La séquence complète de F1 est donnée par SEQ ID NO : 1 : dans la liste de séquence annexée. F1 comprend :

- 488 paires de base sur l'extrémité 3' de l'ADNr 25 S,
- 35 - la région d'espacement intergénique 1 (IGR<sub>1</sub>),
- l'ADNr 5 S

- et la région d'espacement intergénique 2 (IGR<sub>2</sub>).

Comme, cela ressort des Figures 2 et 3, l'IGR<sub>2</sub> (ou NTS = séquence non transcrite) comprend deux séquences sources itératives F $\alpha$  et F $\beta$ . Tandis que l'IGR<sub>1</sub> renferme les deux séquences-sources itératives F $\gamma$ 1 et F $\gamma$ 2.

5

### 1.3.1. F $\alpha$

La séquence-source itérative F $\alpha$  a une longueur de 1256 paires de base. (base 2256 à base 3512). F $\alpha$  est représentée schématiquement sur la Figure 5. La Figure 6 et la SEQ ID NO : 2 : de la liste de séquences annexée, donnent le séquençage complet de F $\alpha$ .

10 F $\alpha$  est constitué de 8 sous-séquences répétées désignées par Kre-0 à Kre-7 respectivement. Les 7 premières sous-séquences répétées Kre-0 à Kre-6 ont une longueur variant entre 164 et 165 paires de base. La dernière sous-séquence Kre-7 comporte 103 paires de bases (voir Figures 5 et 6).

Le fragment d'ADN correspondant à F $\alpha$  à une teneur en bases GC de 35 % molaire.

15

### 1.3.2. F $\beta$

L'autre séquence-source itérative d'un fragment F susceptible d'entrer dans la constitution d'une sonde selon l'invention, est compris dans l'IGR<sub>2</sub> et désigné par la référence F $\beta$ . Le séquençage de F $\beta$  est représenté sur la Figure 7A et sur la SEQ ID :

20 NO : 3 : de la liste de séquences annexée. La séquence-source F $\beta$  comprend 9 sous-séquences répétées comprises entre la base 4823 et la base 4903. Ces sous-séquences ont une longueur variable entre 8 et 10 bases. Elles contiennent un noyau commun de 6 bases constituant un motif commun (référéncé sur la Fig. 7A par Com 1), répété soit directement (les uns à la suite des autres) soit en étant séparés par des

25 espacements de quelques bases nucléotidiques.

### 1.3.3. F $\gamma$ 1 et F $\gamma$ 2

Ces deux séquences-sources itératives F $\gamma$ 1 et F $\gamma$ 2 sont localisées dans l'IGR<sub>1</sub>. F $\gamma$ 1 comprend 9 sous-séquences répétées de longueur variant entre 6 et 16 bases, ces

30 sous-séquences s'étalant de la base 923 à la base 1011 (Figure 7B). F $\gamma$ 2 s'étend de la base 1012 à la base 1151 et comporte 9 sous-séquences de longueur variant entre 10 et 17 bases (Figure 7C). Les sous-séquences de F $\gamma$ 2 sont des copies très proches les unes des autres. Deux noyaux aux motifs communs de 6 et 7 bases peuvent être isolés respectivement dans F $\gamma$ 1 et F $\gamma$ 2. Les motifs communs de F $\gamma$ 1 et F $\gamma$ 2 sont identifiés par

35 Com 2 et Com 3 sur les Figures 7B et 7C respectivement.

A l'instar de Com 1 de Fy, Com 2 et Com 3 sont répétés soit directement (les uns à la suite des autres) soit en étant séparés par des espacements de quelques bases nucléotidiques.

## 5 EXEMPLE 2 : ETUDE DE L'HOMOLOGIE ENTRE LES HUIT SOUS-SEQUENCES DE F $\alpha$

L'homologie des sous-séquences Kre-0 à Kre-7 a été étudié en comportant les sous-séquences après alignement optimal.

Les résultats obtenus sont donnés dans le tableau I qui suit :

10

**TABLEAU I**

*POURCENTAGE DE BASES QUI "ALLUMENT" OU QUI CORRESPONDENT ENTRE LES 8 SOUS-SEQUENCES Kre-0 A Kre-7*

15

% de bases correspondantes (no. des bases différentes)								
Séquences	Kre-0	Kre-1	Kre-2	Kre-3	Kre-4	Kre-5	Kre-6	Kre-7*
kre-0	100							
Kre-1	71 (47)	100						
Kre-2	77(39)	95(9)	100					
Kre-3	77(39)	95(9)	100	100				
Kre-4	77(39)	95(9)	100	100	100			
Kre-5	67(56)	90(17)	93(11)	93(11)	93(11)	100		
Kre-6	71(47)	96(6)	94(10)	94(10)	94(10)	91(12)	100	
Kre-7*	66 (35)	88(12)	96(4)	96(4)	96(4)	97(3)	90(10)	100

\* Les 103 bases de Kre-7 ont été comparées avec les 103 premières bases de Kre-n =

20 0 à 6.

EXEMPLE 3 : HYBRIDATION DE CHROMOSOMES DE *CANDIDA KRUSEI* ET DE *CANDIDA ALBICANS* AVEC DES SONDAS COMPRENANT LES FRAGMENTS F1 OU F $\alpha$  ET F $\beta$  OU F $\gamma$ 1, F $\gamma$ 2

Cet exemple vise à démontrer que la localisation des fragments F1, F $\alpha$ , F $\beta$  F $\gamma$ 1 et F $\gamma$ 2  
5 constituant la sonde selon l'invention est seulement sur les chromosomes contenant des cistrons d'ARNr.

3.1. Les souches mises en oeuvre sont LMCK31 pour *Candida krusei* et la souche 3153A pour *Candida albicans*.

10

3.2. Les fragments F1, F $\alpha$ , F $\beta$  F $\gamma$ 1 et F $\gamma$ 2 mis en oeuvre sont ceux isolés dans les Exemples 1 et 2.

3.3. La technique mise en oeuvre est celle d'électrophorèse transversale à  
15 champ pulsé. Des plaques d'agarose contenant l'ADN chromosomique sont préparées à partir de suspensions de culture de souche *C. krusei* LMCK 31 et *C. albicans* 3153A. Les cellules sont converties en sphéroplastes par la méthode de shareman et al, selon laquelle on met en oeuvre la ZYMOLYASE 20 T. Les sphéroplastes sont tout  
20 d'abord lavés deux fois dans une solution de sorbitol molaire puis remis en suspension à raison de 10<sup>9</sup> cellules/ml. Les sphéroplastes sont mélangées avec un volume = d'agarose seaplaque GT à 1 % (FMC bioproducts, Rockland ME) et immédiatement versé dans un moule. Les plaques sont traitées avec du SARCOSYL à 1% et avec une protéinase K titrant 100 mg/ml et enfin incubé à 50° pendant 60 heures. Les plaques  
25 sont ensuite lavées dans une solution tampon TE (10 mM tris-HCL), pH 8, 1mM EDTA) et stockée dans ce milieu à 4° C. Les ADN chromosomiques sont séparés par électrophorèse sur gel à champ pulsé dans un dispositif Gene line (BECKMAN, FULLERTON, CA) en utilisant de l'agarose LE 0,65 % (FMC bioproducts, Rockland me) ainsi qu'un tampon tris - acétate 0,04 M tris acétate, 0,001 M EDTA pH8). L'électrophorèse est menée selon les cinq étapes telles que décrites ci-après : 65 V  
30 pendant 6 heures avec une durée de commutation d'une minute ; 65 V pendant 12 heures avec une durée de commutation de 2 mn ; 65 V pendant 16 heures avec une durée de commutation de 4 mn ; 65 V pendant 20 heures avec une durée de commutation de 7 mn et 65 V pendant 18 heures avec une durée de commutation de 10 mn. Les gels sont ensuite révélés avec du bromure d'ethidium, photographiés puis  
35 transférés sur une membrane de nylon ZETABIND. L'hybridation et le lavage sont réalisés selon la méthodologie de CHURCH et GILBERT. Les bandes

chromosomiques sont sondées avec les différentes sondes marquées radioactivement : F1, F $\alpha$  et F $\beta$  F $\gamma$ 1 et F $\gamma$ 2, selon la méthode d'hybridation sur membrane.

### 3.4. Résultats

5 La Figure 8A montre les gels d'électrophorèse à champ pulsé qui montre que l'on a séparé 5 chromosomes majeurs pour *Candida krusei* et 7 chromosomes majeurs pour *Candida albicans*.

La Figure 8 B montre que les sondes F1, F $\alpha$  et F $\beta$ , F $\gamma$ 1 et F $\gamma$ 2 hybrident uniquement avec les trois plus gros chromosomes 1, 2, 3 de *Candida krusei*. Les sondes  
10 n'hybrident avec aucun des 7 chromosomes de *Candida albicans*. Ce qui démontre la spécificité inter-espèces desdites sondes. La Figure 8A montre également un minichromosome m (pour *Candida krusei*).

La Figure 8 montre une hybridation sur membrane dans laquelle les sondes ont été mises en contact avec des chromosomes pendant 4 jours.

15 Le code P sur les Figures 8A à 8B correspond à la zone de dépôt de l'échantillon.

#### EXEMPLE 4 : TEST DE DETECTION ET D'IDENTIFICATION ENTRE ESPECES

La Figure 9A montre le gel d'électrophorèse de l'ADN de souches de différentes  
20 espèces de levures, digéré par l'enzyme *EcoRI*. La Figure 9B montre l'hybridation selon Southern correspondante lorsque les sondes F1, F $\alpha$  et F $\beta$ , F $\gamma$ 1 et F $\gamma$ 2 sont utilisées, en étant marquées à la peroxydase. Le protocole expérimental est le même que celui décrit décrit dans l'Exemple 1 supra ainsi que dans l'Exemple 3 du WO-95/11991. Les sondes hybrident uniquement avec l'ADN de *C. krusei* (CBS 573T, bande n° 10) et  
25 pas avec l'ADN des autres levures testées. Ceci démontre la spécificité interespèce desdites séquences.

Dans les essais correspondant aux Fig. 9A et 9B, les microorganismes considérés sont les suivants :

- 30 S : Standard : Phage lambda digéré par *HindIII*  
1 : *C. tropicalis* CBS 94T  
2 : *C. parapsilosis* CBS 604T  
3 : *C. guilliermondii* CBS 6021T  
4 : *C. lusitaniae* CBS 6936T  
35 5 : *C. kefir* CBS 607T  
6 : *C. albicans* serotype B

- 7 : *C. albicans* serotype A  
 8 : *C. Albicans* ATCC 2091  
 9 : *C. valida* CBS 638T  
 10 : *C. Krusei* CBS 573 T  
 5 11 : *Y. lipolytica* CBS 6124 T  
 S : Standard : Phase lambda digéré par *Hind*III

EXEMPLE 5 : TEST DE TYPAGE INFRA-SPECIFIQUE POUR DIFFERENTES SOUCHES DE *C. Krusei*

10 La Figure 10A montre le gel d'électrophorèse de l'ADN de plusieurs souches de l'espèce *C. Krusei*, digéré pour l'enzyme *Eco*RI. La Figure 10B montre l'hybridation selon Southern correspondante lorsque les sondes F1, F $\alpha$  et F $\beta$ , F $\gamma$ 1 et F $\gamma$ 2 sont utilisées, marquées à la peroxydase. Le protocole expérimental est le même que celui  
 15 décrit dans l'Exemple 1 supra. (ou Exemple 3 du WO 95/11991). Les sondes hybrident avec toutes les souches de *C. krusei* testées, ce qui illustre leur intérêt pour la détection et l'identification de cette espèce. En outre, ces sondes révèlent un polymorphisme de la taille des fragments de restriction selon les sondes. Ceci illustre  
 20 l'intérêt de ces sondes pour différencier les souches entre-elles au niveau infraspécifique et donc pour le typage des souches dans le cadre d'études épidémiologiques ou le traçage des souches en milieu industriel. Toujours à titre d'exemple les souches référencées K62-K64 dont les profils d'hybridation sont identiques, proviennent d'un même patient.

25 EXEMPLE 6 : TEST D'IDENTIFICATION ET TYPAGE DE SOUCHES DE *C. KRUSEI* PAR REACTION DE POLYMERISATION EN CHAÎNE (PCR) DE LA RÉGION F $\alpha$  (OU CKRS-1)

La Figure 11 montre le résultat de l'amplification sélective (selon la méthode PCR) de la région F $\alpha$  (CKRS-1) à l'aide des amorces Arno1 :  
 30 CGTAGGATACTAACCACAGC et Arno 2 : GGCCAACACATACATACCTT dont les séquences sont choisies dans le fragment F1 de *C. krusei* (positions 2138-2167 et 3989-3970). Les conditions d'amplification sont celles décrites par Carlotti et al. 1994 : Systematic and Applied Microbiology 17, 380-386. La procédure d'amplification par PCR est la suivante : 50  $\mu$ l de la solution d'ADN à tester (ADN de  
 35 *C. krusei* et ADN d'autres espèces de levures) contenant environ 10 à 100 pg sont



ajoutés à 50 µl du mélange réactionnel contenant la Taq polymérase dans un tube eppendorf et placé sur un thermocycleur. La dénaturation initiale dure 3 minutes à 92° C. Elle est suivie de 30 cycles constitués de : 1 minute à 92°C, 1 minute à 60°C, 2 minutes à 72° C. L'extension finale dure 10 minutes à 72° C.

- 5 Les différentes souches testées dans cet exemple sont désignés par les références R, 12 à 26 sur la Figure 11. La légende de ces références est donné ci-après :

R : Raoul

12 : *C. krusei* CBS 573T

10 13 : *C. krusei* LMCK 31

14 : controle Négatif

15 : *S. cerevisiae* CBS 1171

16 : *C. albicans* ATCC 2091

17 : *C. norvegensis* LM Q243

15 18 : *C. glabrata* CBS 138T

19 : *Y. lipolytica* CBS 6124T

20 : *C. kefir* CBS 607T

21 : *C. lusitaniae* CBS 6936T

22 : *C. guilliermondii* CBS 6021T

20 23 : *C. parapsilosis* CBS 604T

24 : *C. rugosa* SIPH

25 : *C. valida* CBS 638T

26 : *C. inconspicua* CBS 180T

- 25 Du fait de la spécificité de la région (CKRS-1) cible et des amorces choisies, seuls les ADNs des souches de *C. krusei* (réf. 12, 13) ont donné un signal d'amplification positif (Figure 11). Les ADNs des autres espèces de levures testées n'ont pas été amplifiés. En outre, les différences de la taille des fragments amplifiés chez les deux souches de *C. krusei* testées traduisent le polymorphisme de la région Fα(CKRS-1).
- 30 Cet exemple illustre bien l'intérêt de tests d'amplification par PCR de la région Fα (CKRS-1) pour l'identification des souches de *C. krusei* et pour leur typage simultané. Cette méthodologie présente l'avantage d'être très rapide et utilisable directement sur les prélèvements biologiques (crachats, sang, selles, LCR) ou alimentaires.

EXEMPLE 7 : TEST DE DETECTION, D'IDENTIFICATION-TYPAGE DES LEVURES DE L'ESPECE *Geotrichum candidum*

- 7.1. L'isolement et le clonage du Fragment  $F_{10}$  comprenant  $F\delta$  compris dans la sonde  $A_2O_{27}$  développée pour *Geotrichum candidum*, ont été réalisés conformément au protocole décrit à l'exemple 1 supra.

La Figure 12 donne la carte de restriction du fragment  $F_{10}$  comprenant la séquence-source itérative  $F\delta$  (GCRS-1).

Abréviations : EV = *EcoRV*, SII = *SacII*, EI = *EcoRI*.

10

7.2. Test de détection/identification

- La Figure 13 montre les résultats d'hybridation, selon la méthode de Southern, de l'ADN de souches des espèces *Geotrichum candidum* et *Geotrichum fermentans* (digérés par *MspI*) avec la sonde  $A_2O_{27}$  spécifique de *Geotrichum candidum*. Les fragments de restriction par *MspI* de l'ADN des souches étudiées ont été séparés par électrophorèse en gel d'agarose à 0,8 % puis transférés sur membrane en nylon. La membrane a été sondée avec environ 100 ng de la sonde  $A_2O_{27}$ . Cette sonde n'hybride qu'avec l'ADN des souches de *Geotrichum candidum* et révèle un polymorphisme de la taille des fragments de restriction.

- 20 La sonde  $A_2O_{27}$  a été développée selon le modèle de la sonde  $F = F1, F\alpha, F\beta, F\gamma1, F\gamma2$ . La sonde  $A_2O_{27}$  est constituée d'une partie  $F_{10}$  dont  $F\delta$  de l'IGR<sub>1</sub> d'une souche de *Geotrichum candidum*. La sonde  $A_2O_{27}$  hybride uniquement avec l'ADN des souches de *Geotrichum candidum* (Fig. 13). Elle n'hybride pas avec l'ADN des souches de *Geotrichum fermentans*, bien que ces deux espèces soient proches. Cet exemple montre bien la spécificité de cette sonde selon l'invention. De plus, cette sonde révèle également un polymorphisme de la taille des fragments de restriction selon les souches de *Geotrichum candidum*. Ceci permet de les différencier entre-elles et donc de les typer.

- 25 Cet exemple illustre la validité du modèle présenté de sonde d'identification-typage pour les levures d'intérêt médical et/ou d'intérêt industriel.

- 30 Les souches référencées sur la Figure 13, N° 27 à 37 correspondant à :

S = Standard

*G. fermentans* CBS 50.57

*G. fermentans* CBS 25.29

*G. candidum* F

5 *G. candidum* A

*G. candidum* B

*G. candidum* LMGC3

*G. fermentans*

*G. candidum* 149.26

10 *G. candidum* LMGC8

*G. candidum* LMGC7

*G. candidum* CBS 557.83

Standard

15

## LISTE DE SEQUENCES

## (1) INFORMATIONS GENERALES:

## (i) DEPOSANT:

- (A) NOM: Université Claude Bernard Lyon I
- (B) RUE: 43, Boulevard du 11 Novembre 1918
- (C) VILLE: Villeurbanne
- (E) PAYS: France
- (F) CODE POSTAL: 69622 Cedex

(ii) TITRE DE L' INVENTION: Sonde et procédé notamment pour la détection et/ou l'identification inter/infraspécifique de levures, de même que les outils génétiques correspondants

(iii) NOMBRE DE SEQUENCES: 5

## (iv) FORME DECHIFFRABLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: Floppy disk
- (B) ORDINATEUR: IBM PC compatible
- (C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS
- (D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

## (vi) DONNEES DE LA DEMANDE ANTERIEURE:

- (A) NUMERO DE LA DEMANDE: FR 96 03835
- (B) DATE DE DEPOT: 22-MAR-1996

## (2) INFORMATIONS POUR LA SEQ ID NO: 1:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 5261 paires de bases
- (B) TYPE: nuclotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(iv) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

GAATTCGGTA AGCGTTGGAT TGITCACCCA CTAATAGGGA ACGTGAGCTG GGTITAGACC	60
GTCGTGAGAC AGGTTAGTTT TACCTACTG ATGAATGTTG TCGCAATAGT AATTGAACIT	120
AGTACGAGAG GAACCGTTCA TTCAGATAAT TGGTTTTTGC GGCTGTCTGA GCAGACACTG	180

CCGCGACGCT ACCATCTGCT GGATAATGGC TGAACGCCTC TAAGTCAGAA TCCATGCTAG	240
AACGCGACGA TTACCTGCCC TCGCACATTT GAGAAGGATA CGAATAAGGC CCTGTGGCCG	300
CAGAACCGTA GCAGGCCGGC AGCGGTGCGC ATGGCGGAAA GGCCGTGTGT GCTTGCCGGC	360
GGATGGCAAT GTCAGGATGC GCTCAGATAA ATCCTATGCA TACGACTTAG ATGTACAACG	420
GGGTATTGTA AGCAGTAGAG TAGCCTTGTT GTTACGATCT GCTGAGATTA AGCCTCTGTT	480
GTCCGATTTG TTGCTTCGGC GCGGAAGGAG CGGTCCGTAG AGACGGGCTG TTGTTTCGGG	540
AGTGTGGGGA GTGTGGGGG ACTTAGGAGC GGGGGTAGTG TGGGGGAGTG TGGGGACTTG	600
GGAGCGGGNG TAGTAGGTGG TACTGATGGT ACGGGGAGGA GCCTACCCTT TTTTGTGCTT	660
TTGGCAAAGA CATTGCAGGA TTTACAAAGT GATTTACGAT ATGAAGTAGG ACTATGTTGT	720
GTGGAGGGGG TGGGGCTGTA GAGAGGATGG ATTTGACGGG CATGACGGGG ACGGCTTGGA	780
GGCCGCCCAT GTCGTATGCC CCGGGGTAGG GTATAGGCCG GGAGGGGCCA GTTGGGAAAA	840
TAGCGAAGAG GGGGGGAGGA GGCTAAGAAG GCGGTGGGGG AGGGGGTAGG TTGATAAGTA	900
GGGGGTGTAG GAGCGTGGT GGAGTGGGAG TGGGAGTGGG TGGTACAGAG TGGGTGTTAC	960
AGAGTGGGTG GTACAGAGTA CAGAGTGGGG AGAGTACAGA GTGGGAGTGG TAGGAGGAAG	1020
AGACAGTACA CAGTAGGAGG AAGAGAGTGG TAGGAAGAGA GTGGTAGGAG GAAGAGCGGC	1080
AGGAGGAAGA GAGTGGTAGG AGGAAGAGTG GTAGGAGGAA GAGTGGTAGG AAGAGAGTGG	1140
TAGGAGGAAG ATGGAATCCT AGACCTTGTA AATTGAGACA AAAGTCCGTC AAAAGTAGA	1200
CAATGGGGGC AGTGGGGGAT TTTCGACAAC AAAAACAAG GGAGGTAAAC AAAGGCAAAAC	1260
AAAGGTAAAC ACAAGAAAA CAGGCAGCAC AGGCAGNAAA AAGATTGCAG CACCTGAGTT	1320
TCGCATATGG TCTCCCACTA CACTACTCGG TCAGGCTCTT AGCAGCTTAA CTACAGTTGA	1380
TCGGACGGTA AACGGTGCTT TCTGCTAGAT ATGGCCGCAA CCGAAAGTAT AGAACGGGCG	1440
GCGTATGGGA GGTTTGATTT TGGGCGGGCA GATTTGACG GGTGCGATAA TCTCGACGGG	1500
ACGCATGTTT GCAGGGCCAT AGCAGACGAA CCAGAGCAGC TGAACAGGCG AGTCAACCAC	1560
CGTTGGAAG AGCGCCGAGA AGTATCTAGT CCCCCCATC GGTGTGGCCA TAGGACGCAA	1620
TCGGCAACAT GGCCAAAATC AAAAAACAAA AATTCCCACT TTTACAAAAT ATATATAATT	1680
TACAAAAAAT AGGTCTGGCA AAAATGCAAT TTTGGAAAAA AATCGAGTAA AAAGTAAATA	1740
TTATGGGCGG AACGGATAGC CGTCGACCGG ACGATGGTGA AAACCGTGGT TGGAAAGCCG	1800
ATATCGGGCC TGGGCGGAGA AAACGGCGAA AAGACAATC CGAACACGAC ATTTCGGCAA	1860
TGTTGCTGTT TCGGCAAAGT CCGAGGCGAA GTCCGAGGCG TAACTCCGAT GGGCGGAGAT	1920
AAAGTGGCGA CGTGAAGTGT GCGAGGGGTG GGGCGTTCTG GAGCGACGCA AAAAAAAAAA	1980
CACGCACAAA AAAAATGGAC GTCGAGAAAA AAAAATTTAC ACGAGCGAGC ACTAAGGTGG	2040
GAATGGACGT CCCCAGAAAT GTCTTCCACT TTGCCACCAC TTTTCTTTT GCACACTTCT	2100
GCTGTATGTC TGTCTGCACC ACCTCTTCCA CAACCAAGCA CACCACAGGT AGGATACTAA	2160
CCACAGCAGG GCTGTCTCTG CAGCCAACCA CCCAGCAGTT TCTTTTGATT TTGATTTTTT	2220
TGATTTTNNT GATTTTTATT TTGACTGATT TTGCTTTTCC CTTGTGTAAA AGTCTCGTCG	2280
ACATTGCAGC TGGATGCAAT GGGCGTTTCT GAACACCAAT CCACTTAGTT GGGTTGATTT	2340
TTCTTACTGT CTGATGAAGA TATGATTGTT TTTTAAAATT CCATTTTTTT TGATTTTATA	2400
TAACTTCACT GTGGGGTGGT TTCCCTTGGC CTGCAAGCTC ATACACATTA CAGTTCCATC	2460
CATTGCACGT TTCTAAACAC TACTCCACTT AGTTGGGCTG TTTATTTCCA CATGCTGATG	2520
AAATGGTTGC TGTTTTTCAG AATTTTCAGTT TTTTGATTTT ATATAACTTC ACTTTGGAGA	2580
TGATTTCCCT TGGACTTTAC AATCACACAC ATTACAGTTC CATCCATTGC ACGTTTCTAA	2640
ACACTACTCC ACTTAGTTGG GCTATTTATT TCCACATGCT GATGAAATGG TTGCTGTTTT	2700
TCAGAATTTT AGTTTTTTTG ATTTTATATA ACTTCACTTT GGAGATGATT TCCCTTGGAC	2760
TTTACAATCA CACACATTAC AGTTCCATCC ATTGCACGTT TCTAAACACT ACTCCACTTA	2820
GTTGGGCTAT TTATTTCCAC ATGCTGATGA AATGGTTGCT GTTTTTTCTA ATTTTCAGTTT	2880
TTTTGATTTT ATATAACTTC ACTTTGGAGA TGATTTCCCT TGGACTTTAC AATCACACAC	2940

ATTACAGTTC CATCCATTGC ACGTTTCTAA AACTACTCC ACTTAGTTGG GCTATTTATT	3000
TCCACATGCT GATGAAATGG TTGCTGTTTT TCAGAAATTC AGTTTTTTTG ATTTTATATA	3060
ACTTCACTTT GGAGATGATT TCCCTTGGAC TTACAATCA CACACATTAC AGTTCCATCC	3120
ATTGCACGTT TCTAAACACT ACTCCACTTA GTTGGGCTGT TTATTTCCAC AGGTTGATGA	3180
AATGGTTGCT GTTTTTCAGA ATTTCACTTT TTACAATTTT AACTAGTTTC ACTTTGGAGA	3240
TGATTTCCCT TGGCATGGAA GCTCACACAC ATTACAGTTC CATCCATTGC ACGTTTCTAA	3300
AACTACTCC ACTTAGTTGG GCTGTTTATT TCCACAGGTT GATGAAATGG TTGCTGTTTT	3360
TCAGAAATTC AGTTTTTTTG ATTTTATATA ACTTCACTTT GGAGATGATT TCCCTTGGAC	3420
TTTACAATCA CACACATTAC AGTTCCATCA ATTACAGTT TCTAAACACT ACTCCACTTA	3480
GTTGGGCTAT TTATTTCCAC AGGTTGATGA AGTTGTCAAG TTTTACAATT TTGGTTGCTG	3540
AAGGTTCCGGC AATTCAGTGT TGTATAATTG AACGACAGAC ACTTCAATCA GCACTGTGCT	3600
TCCCTTTTAT GTATTCTCCA ACGTAGTATC TTGACAAAGT ATATATATGA CTGCCACTTT	3660
GATTATTTTG ATAGCTTTGT ATTTGACCAC TAATTTTGTA TTTGACCACC AATTTGAAAA	3720
TGCCAATACC AATTTGCATT TGCCAGGGAA TTTCAATACC AATTTGCATT GCCAGGGAAT	3780
TTCAATACCA ATTTCCAGGA ATTTCAATAC TAATTACCAC CATTTCAGG AATTAATTTT	3840
GTCGAAAAAT TAAAAAGNC GTCGAGAAAT TTTGTGTTTG TTTTGCAAAC AAAATATCAC	3900
ACACAACACT GAATACACGG AAGTTCTGTT GGACACAGAG CCCCACACAC ACAACGTACA	3960
TGCAATTGGA AGGTATGTAT GTGTTGGCCA GGAGCAGGCG CAACTGTGCC GTGCTGCTTG	4020
TGCATAATAC AGCGGCGTAC CGACAGAAAT AGAGCCCCC AAAATTTGAT GGAAAAATAG	4080
TGGACAGTTG GACTTTGCAA GTTGGACGCA GTGGAATAGA CAACCAAACA CCCCATTAA	4140
CAAGCCGATG GCCGATTGAC CCAATGCCAA TCAGTATGGA TCGATAACTG TCCAATGAA	4200
CGAGTAACAA GAAGATGTTT TCCAAATGTC AGAATTTCAA ATCGGATGTT CCACTGATGG	4260
TCTGGGCGTT GCAGTCTGGG CTCCAGGTGT GTCAACTGGA CATTGCCAAG CGTGCACCGA	4320
TGTTGCAAGA CGATTACCAC GGCAACAAAC CGGTGTTTAC GACAAGAACA CTGGCAGATG	4380
CATCGAATAC TGATTTGGTG CAATTTCAAG GTTTGGGAGT TTTGTATAAT GTGTGTATAT	4440
TATGATGCGA CAACATGCGC CGGTGTGTAA ATAGTGTATT GCAACCTGTA AATAAACTAA	4500
ATCGACAAGG AATTGGTGGA ATTAGACGTA AAATAGTGTT GCGGAAAATA CGTTTTTTGT	4560
TTCTGGGGGA CCTATGACGA TACTTTTTGG ACGCACATGT GGAGCGGCAT GAAATTAGTG	4620
TGTGGCCGAG TTTGGGAGGA TATTCCCCAG GCGGTTGCCA AGATATTTGA AATAATCTGA	4680
ATTTCAAAAA TGTCGGTGCT CTTTCGAAAT GCCATGTCCA GAGGATTTGA AATTTTGGGA	4740
TGGAGCTGTG TTGGCTGGGT TGATGTTAAG GGAATGATAT ATCTGCGCTT TTCAATGTAT	4800
TATGTTTCAG ATTTTITGAG ATTTTGAAT CTTTGAATTT TTTGAAATCT TTTGAATTTT	4860
TTGAATCTTT TGAATTTTTT GAAATCTTTT GAATTTTTTT CAAAGTCTAT AAATGTTATT	4920
GTGTTTTTGG TTTTATTTTA GCTTTTGCAA AATGTTATTT CATTATACAG TTCAAATAA	4980
TTTAATTGCC TATATGTTTT ACAATATAAT ATATTTTATT TCAGATTTCA AAATAATTTT	5040
TTCCATTAAA CGTTTATAAT ACATTATTTT ACTTCAGATT TCAGGATATA CTATTCATTC	5100
CCACACTTTA AAATATATTA TTCAATTTCA CACACCAAAG TGTAATTTT AACTTCACAT	5160
TTCAAATAA AATATTTGAC AATACCTAAA TATGGCATTT TTATTTCCGA TTTCGTTAGA	5220
TATTCCTTAG TTATACAATC CCAAACATGC TATTCGAATT C	5261

## (2) INFORMATIONS POUR LA SEQ ID NO: 2:

## (1) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1257 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple

## (D) CONFIGURATION:- linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(iv) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

TTTCCCTTGT GTAAAAGTCT CGTCGACATT GCAGCTGGAT GCAATGGGCG TTTCTGAACA	60
CCAATCCACT TAGTTGGGTT GATTTTTCTT ACTGTCTGAT GAAGATATGA TTGTTTTTTA	120
AAATTCCATT TTTTGTGATT TTATATAACT TCACTGTGGG GTGGTTTCCC TTGGCCTGCA	180
AGCTCATACA CATTACAGTT CCATCCATTG CACGTTTCTA AACACTACTC CACTTAGTTG	240
GGCTGTTTAT TTCCACATGC TGATGAAATG GTTGCTGTTT TTCAGAATTT CAGTTTTTTG	300
ATTTTATATA ACTTCACTTT GGAGATGATT TCCCTTGGAC TTTACAATCA CACACATTAC	360
AGTTCCATCC ATTGCACGTT TCTAAACACT ACTCCACTTA GTTGGGCTAT TTATTTCCAC	420
ATGCTGATGA AATGGTTGCT GTTTTTCAGA ATTTTCAGTTT TTTTGATTTT ATATAACTTC	480
ACTTTGGAGA TGATTTCCCT TGGACTTTAC AATCACACAC ATTACAGTTC CATCCATTGC	540
ACGTTTCTAA ACACTACTCC ACTTAGTTGG GCTATTTATT TCCACATGCT GATGAAATGG	600
TTGCTGTTTT TCAGAATTTT AGTTTTTTTG ATTTTATATA ACTTCACTTT GGAGATGATT	660
TCCCTTGGAC TTTACAATCA CACACATTAC AGTTCCATCC ATTGCACGTT TCTAAACACT	720
ACTCCACTTA GTTGGGCTAT TTATTTCCAC ATGCTGATGA AATGGTTGCT GTTTTTCAGA	780
ATTTTCAGTTT TTTTGATTTT ATATAACTTC ACTTTGGAGA TGATTTCCCT TGGACTTTAC	840
AATCACACAC ATTACAGTTC CATCCATTGC ACGTTTCTAA ACACTACTCC ACTTAGTTGG	900
GCTGTTTATT TCCACAGGTT GATGAAATGG TTGCTGTTTT TCAGAATTTT AGTTTTTACA	960
ATTTCAACTA GTTTCACTTT GGAGATGATT TCCCTTGGCA TGGAAGCTCA CACACATTAC	1020
AGTTCCATCC ATTGCACGTT TCTAAACACT ACTCCACTTA GTTGGGCTGT TTATTTCCAC	1080
AGGTTGATGA AATGGTTGCT GTTTTTCAGA ATTTTCAGTTT TTTTGATTTT ATATAACTTC	1140
ACTTTGGAGA TGATTTCCCT TGGACTTTAC AATCACACAC ATTACAGTTC CATCAATTAC	1200
ACGTTTCTAA ACACTACTCC ACTTAGTTGG GCTATTTATT TCCACAGGTT GATGAAG	1257

(2) INFORMATIONS POUR LA SEQ ID NO: 3:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 81 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(iv) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

TTTTGAATCT TTGAATTTTT TGAAATCTTT TGAATTTTTT GAATCTTTTG AATTTTTTGA 60  
AATCTTTTGA ATTTTTTTCA A 81

(2) INFORMATIONS POUR LA SEQ ID NO: 4:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 89 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(iv) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

AGTGGGAGTG GGAGTGGGTG GTACAGAGTG GGTGGTACAG AGTGGGTGGT ACAGAGTACA 60  
GAGTGGGGAG AGTACAGAGT GGGAGTGGT 89

(2) INFORMATIONS POUR LA SEQ ID NO: 5:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 140 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADN (génomique)

(iii) HYPOTHETIQUE: NON

(iv) ANTI-SENS: NON

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 5:

AGGAGGAAGA GACAGTACAC AGTAGGAGGA AGAGAGTGGT AGGAAGAGAG TGGTAGGAGG 60  
AAGAGCGGCA GGAGGAAGAG AGTGGTAGGA GGAAGAGTGG TAGGAGGAAG AGTGGTAGGA 120  
ACAGAGTGGT AGGAGGAAGA 140



**REVENDEICATIONS :**

1 - Sonde, notamment, de détection spécifique et/ou intraspécifique de levures, du type de celle choisie parmi les outils génétiques (ou apparentés) suivants :

- au moins une partie :

- 5                   \*       d'au moins un fragment F d'acide nucléique ADN et/ou  
                          ARN recopié, construit et/ou isolé à partir d'au moins  
                          une séquence-source d'ADN ribosomal (ADNr) :  
                          →       ne codant pas pour la synthèse d'ARN ribosomal  
                                      (ARNr),  
10                   →       et localisée dans la région d'espacement  
                          intergénique (IGR) non transcrite (NTS) et  
                          comprise entre l'ADNr codant pour la sous-unité  
                          25 S d'ARNr, et l'ADNr codant pour la sous-  
                          unité 18 S d'ARNr,

15                   \*       et/ou d'au moins un analogue Fa de ce fragment,  
                          résultant de la dégénérescence du code génétique ;

                          \*       et/ou d'au moins un fragment Fc d'ADNc  
                          complémentaire du fragment F,

20                   -       au moins une partie des produits de transcription primaire du  
                          fragment F et/ou Fa et/ou Fc.

                          -       et une association des outils énumérés ci-dessus.

                          caractérisé en ce que la séquence-source d'ADNr pour F et/ou Fa  
                          et/ou Fc est sélectionnée de telle sorte

- 25                   ➤       qu'elle soit itérative  
                          ➤       qu'elle comprenne au moins 2, de préférence de  
                                      2 à 12 sous-séquences répétées,  
                          ➤       et que le degré d'homologie entre ces sous-  
                                      séquences soit supérieur ou égal à 50 %, de  
                                      préférence 55 %, et plus préférentiellement  
30                   encore à 60 %.

2 - Sonde selon la revendication 1, caractérisée en ce que la séquence-source de F est localisée dans la région d'espacement intergénique 2 (IGR<sub>2</sub>) non transcrite (NTS) et comprise entre l'ADNr codant pour la sous-unité 5S d'ARNr et l'ADNr codant pour la sous-unité 18 S d'ARNr.

35                   3 - Sonde selon la revendication 1 ou 2, caractérisée en ce que la séquence-source itérative de F correspond à une séquence F<sub>α</sub> et comprend au moins trois sous-

séquences répétées comportant chacune au moins 80, de préférence au moins 120 et plus préférentiellement encore entre 150 et 300 bases nucléotidiques.

- 4 - Sonde selon l'une quelconque des revendications 1 à 3, caractérisée en ce que la séquence-source itérative de F comprend au moins une sous-séquence de longueur tronquée par rapport aux autres sous-séquences, cette sous-séquence tronquée étant de préférence terminale.

- 5 - Sonde selon la revendication 1 ou 2, caractérisée en ce que la séquence-source itérative de F correspond à une séquence  $F\beta$  et comprend au moins trois sous-séquences répétées comportant chacune au moins 4, de préférence au moins 6 et, plus préférentiellement encore entre 7 et 25 bases nucléotidiques ; chaque sous-séquence comprenant au moins un motif commun (Com 1) à toutes les sous-séquences et ayant une longueur inférieure ou égale à celle de la sous-séquence à laquelle il appartient.

- 6 - Sonde selon la revendication 1, caractérisée en ce que la séquence-source de F :

- 15 • est localisée dans la région d'espacement intergénique 1 (IGR<sub>1</sub>) non transcrite (NTS) et comprise entre l'ADNr codant pour la sous-unité 5S d'ARNr et l'ADNr codant pour la sous-unité 25 S d'ARNr

- et correspond à au moins une séquence  $F\gamma$  et comprend au moins trois sous-séquences répétées comportant chacune au moins 4, de préférence au moins 5 et, plus préférentiellement encore entre 5 et 50, voire entre 5 et 45, et mieux encore entre 5 et 40 bases nucléotidiques, chaque sous-séquence comprenant au moins un motif commun (Com 2 ou 3) à toutes les sous-séquences et ayant une longueur inférieure ou égale à celle de la sous-séquence à laquelle il appartient.

- 25 un cas de figure particulièrement préféré étant celui dans lequel F correspond à au moins deux séquences  $F\gamma$  :  $F\gamma_1$  et  $F\gamma_2$  ayant respectivement avantageusement de 5 à 20 et de 9 à 25 bases nucléotidiques.

- 7 - Séquence d'ADN caractérisée en ce qu'elle est choisie dans le groupe des séquences annexées suivantes : SEQ ID NO : 1 ; ; SEQ ID NO : 2 ; ; SEQ ID NO : 3 ; ; SEQ ID NO : 4 ; ; SEQ ID NO : 5 ; ; les SEQ ID NO : 2 : à : 5 : étant particulièrement préférés.

- 8 - Sonde selon au moins l'une des revendications 1 à 6, caractérisée en ce que :

- 35  $F\alpha$  correspond à SEQ ID NO : 2 :  
 $F\beta$  correspond à SEQ ID NO : 3 :  
 $F\gamma_1$  correspond à SEQ ID NO : 4 :  
 $F\gamma_2$  correspond à SEQ ID NO : 5 :

- 9 - Sonde selon l'une quelconque des revendications 1 à 8, caractérisée :

- en ce qu'elle comprend au moins une partie d'au moins l'un des fragments  $F\alpha$ ,  $F\beta$ ,  $F\gamma_1$ ,  $F\gamma_2$ , le fragment  $F\alpha$  étant particulièrement préféré.
- en ce qu'elle est appliquée :
  - ▷ à la détection/identification spécifique et infraspécifique de levures,
  - ▷ ou comme marqueur épidémiologique,
  - ▷ ou encore comme traceur de souche,
  - ▷ ou bien enfin dans le cadre d'une stratégie thérapeutique visant des affections à levures.

10       **10 - Procédé de détection et d'identification de levures caractérisé en ce qu'il consiste à mettre en oeuvre au moins une sonde selon l'une quelconque des revendications 1 à 6, 8 et 9.**

15       **11 - Procédé selon la revendication 10, caractérisé en ce que :**

- on extrait l'ADN total génomique des souches à étudier,
- on soumet éventuellement cet ADN total à une digestion enzymatique à l'aide d'au moins une enzyme de restriction,
- on dénature l'ADN total éventuellement digéré,
- on met en présence l'ADN total ainsi dénaturé, avec la sonde dotée d'au moins un marqueur, de façon à réaliser l'hybridation,
- on élimine l'ADN et la sonde non hybridée,
- et on révèle l'hybridation à l'aide du marqueur.

25       **12 - Amorce oligonucléotidique, utilisable notamment pour la réaction de polymérisation en chaîne (PCR), caractérisée en ce qu'elle est constituée par la sonde selon l'une quelconque des revendications 1 à 8.**

30       **13 - Réactif de détection/identification de levures comprenant au moins une sonde selon l'une quelconque des revendications 1 à 8.**

30       **14 - Application de la sonde selon l'une quelconque des revendications 1 à 6, 8 et 9, pour la détection, l'identification inter-espèces et/ou la différenciation infraspécifique des levures notamment des espèces *Candida krusei* ou *Geotrichum candidum*.**

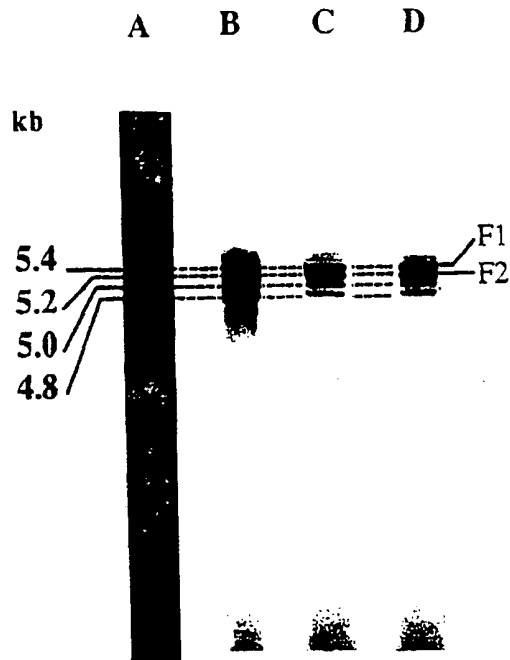
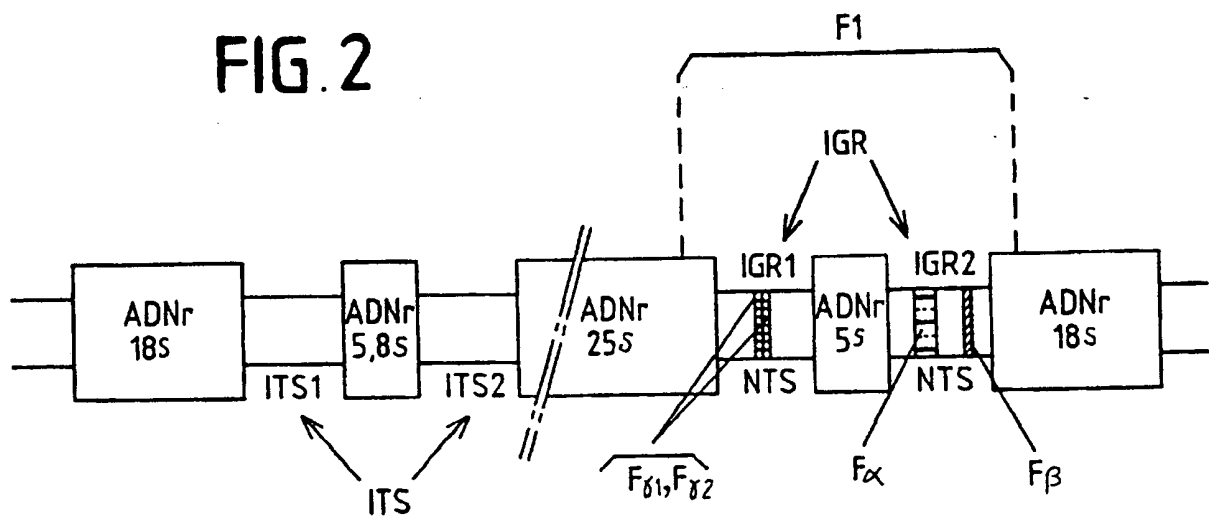


FIG.1



2/9

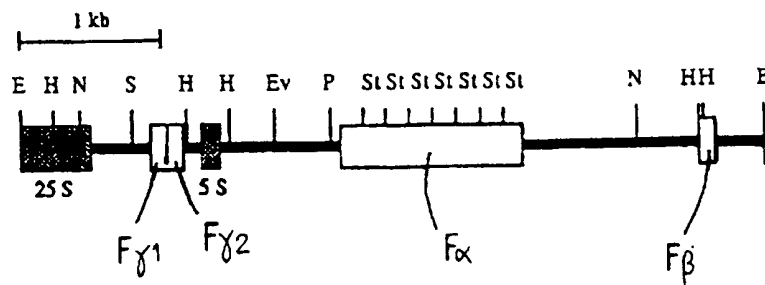


FIG. 3

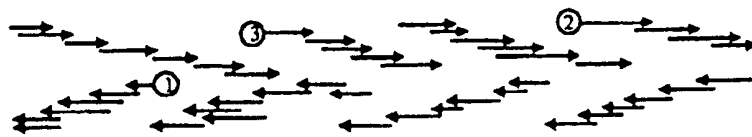


FIG. 4

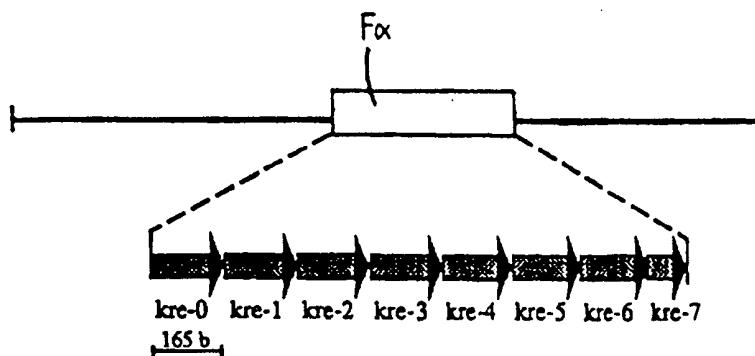


FIG. 5

FEUILLE DE REMPLACEMENT (REGLE 26)

3/9

		«ivrl»	
kre-0	1 *****TGTAAG*GT***G*CG*****G***C*GG* <u>TGCAATGGG</u> **	»ivrl»	50
	2256^		
kre-1	1 TTTCCTTGGCCTGCAAGCTCATAACATTACAGTTCCATGGATTGGACG		50
kre-2	1 *****A**TT*CAA***C*****		50
kre-3	1 *****A**TT*CAA***C*****		50
kre-4	1 *****A**TT*CAA***C*****		50
kre-5	1 *****A**TT*CAA***C*****		50
kre-6	1 *****A**G*****C*****		50
kre-7	1 *****A**TT*CAA***C*****		50
kre-0	51 *****G*****C*A*****T**A**T**CTT**TGT*****		100
kre-1	51 TTTCTAAACACTACTCCACTTAGTTGGGCTGTTTATTTCCACATGCTGAT		100
kre-2	51 *****A*****		100
kre-3	51 *****A*****		100
kre-4	51 *****A*****		100
kre-5	51 *****G*T***		100
kre-6	51 *****G*T***		100
kre-7	51 *****A*****G*T***		100
		<<pal1>>	
kre-0	101 ***GATA*GAT*****TAA***C**T*****T*** <u>TTATATAA</u> **		150
kre-1	101 GAAATGGTTGCTGTTTTCAGAATTTTCAGTTT-TATTATATAACT		149
kre-2	101 *****T*****		150
kre-3	101 *****T*****		150
kre-4	101 *****T*****		150
kre-5	101 *****ACA***C*AC*GT*		150
kre-6	101 *****T*****		150
kre-7	101 ***.....		103
	^3512		
kre-0	151 *****G***-G**G.....		164
kre-1	151 TCACTTTGGAGATGA.....		164
kre-2	151 *****		165
kre-3	151 *****		165
kre-4	151 *****		165
kre-5	151 *****		165
kre-6	151 *****		165
kre-7	151 .....		103

FIG. 6

4/9

N° de sous-séquences	Base de départ	Bases alignées dans les sous-séquences	Longueur des sous-séquences
	4823 ↘	<u>Com 1</u>	
1	(4823)	TTTGAATC	9
2	(4832)	-TTTGAATTT	9
3	(4841)	-TTTGAATC	9
4	(4850)	-TTTGAATTT	9
5	(4859)	-TTTGAATC	8
6	(4867)	TTTGAATTT	10
7	(4877)	-TTTGAATC	9
8	(4886)	TTTGAATTT	10
9	(4896)	TTTCAA	8
		^4903	

FIG. 7A

	923 ↘	<u>Com 2</u>	
1	(923)	AGTGGG	6
2	(929)	AGTGGG	6
3	(935)	AGTGGG - TGGTACAG	14
4	(949)	AGTGGG - TGGTACAG	14
5	(963)	AGTGGG - TGGTACAG	14
6	(977)	AGT - - - - - ACAG	7
7	(984)	AGTGGGAGAGTACAG	16
8	(1000)	AGTGGG	6
9	(1006)	AGTGGT	6
		^1011	

FIG. 7B

	1012 ↘	<u>Com 3</u>	
1	(1012)	AGGAGGAGGACAGTACACAGT	23
2	(1035)	AGGAGGAGGAGAGTGGT	17
3	(1052)	---AGGAGGAGAGTGGT	14
4	(1066)	AGGAGGAGGAGCGGC	15
5	(1081)	AGGAGGAGGAGAGTGGT	17
6	(1098)	AGGAGGAGGAGAGTGGT	15
7	(1113)	AGGAGGAGGAGAGTGGT	15
8	(1128)	---AGGAGGAGAGTGGT	14
9	(1142)	AGGAGGAGGAG	10
		^1151	

FIG. 7C

FEUILLE DE REMPLACEMENT (REGLE 26)

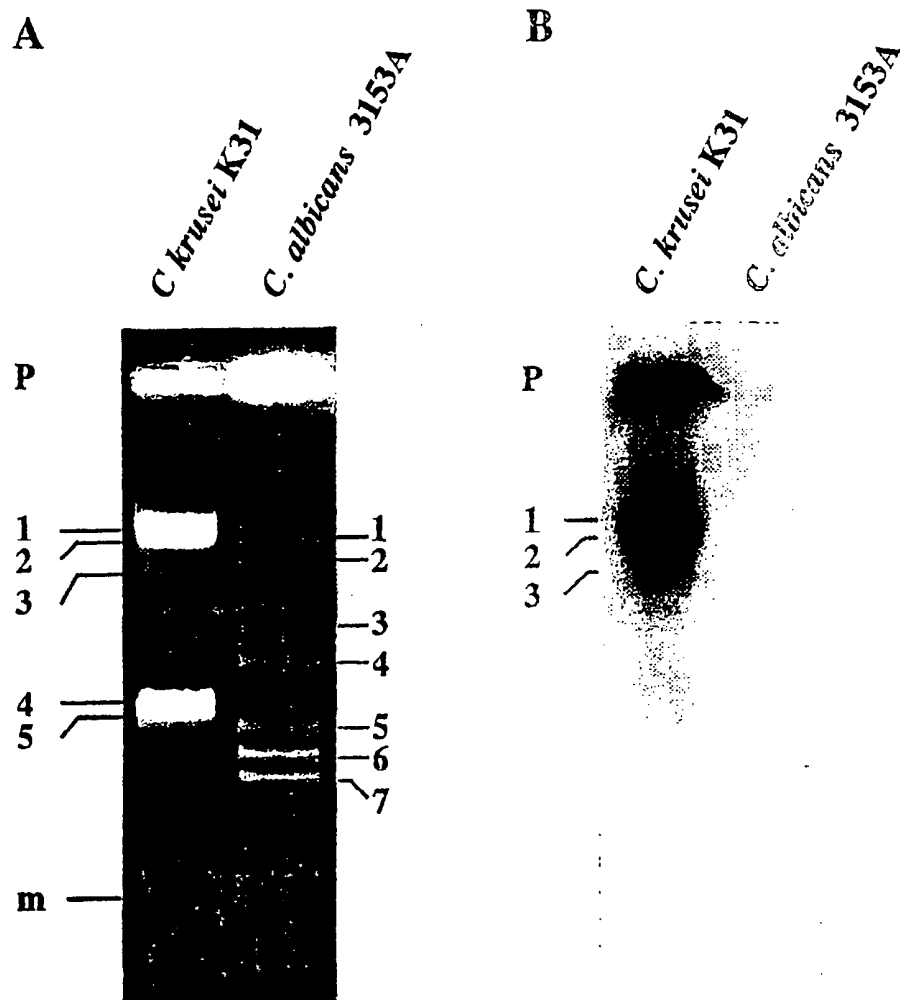


FIG.8A

FIG.8B



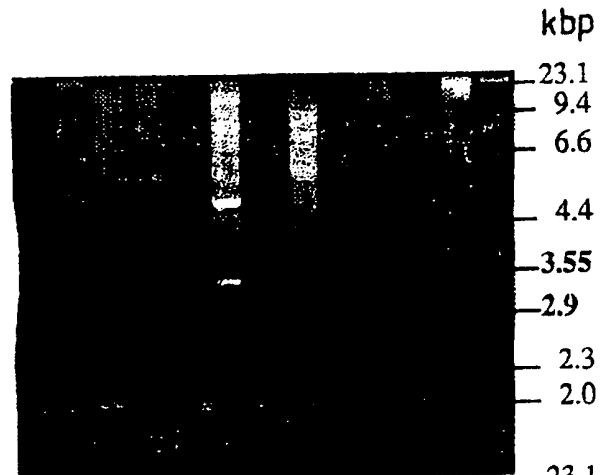


FIG.9A



FIG.9B

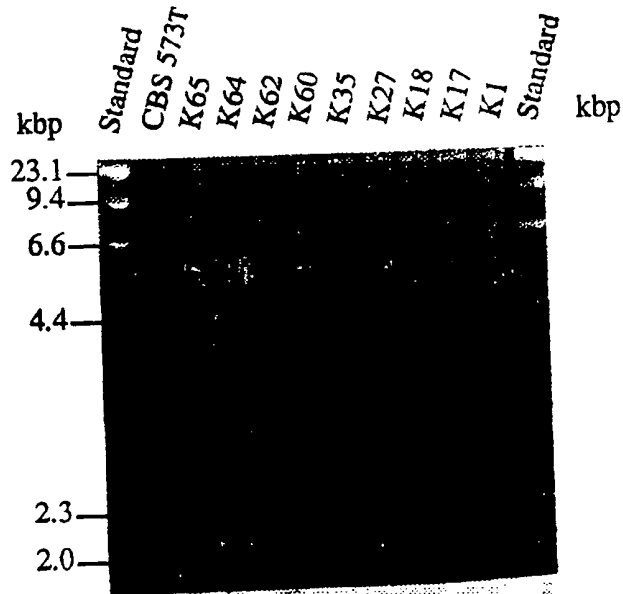


FIG.10A

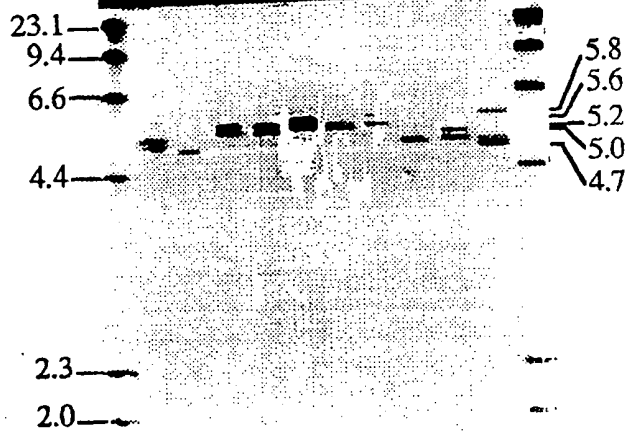


FIG.10B

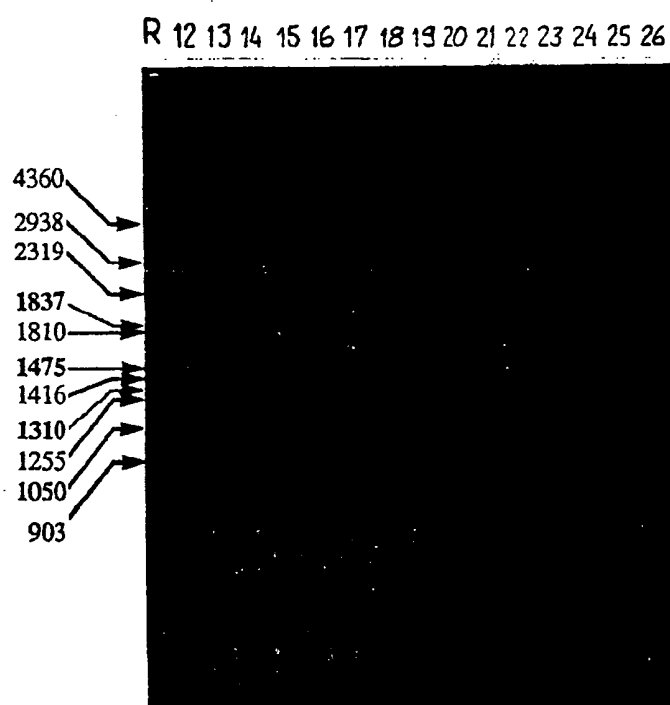


FIG.11

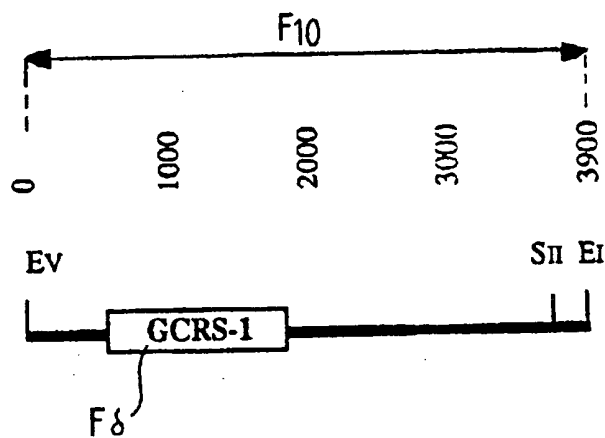


FIG.12

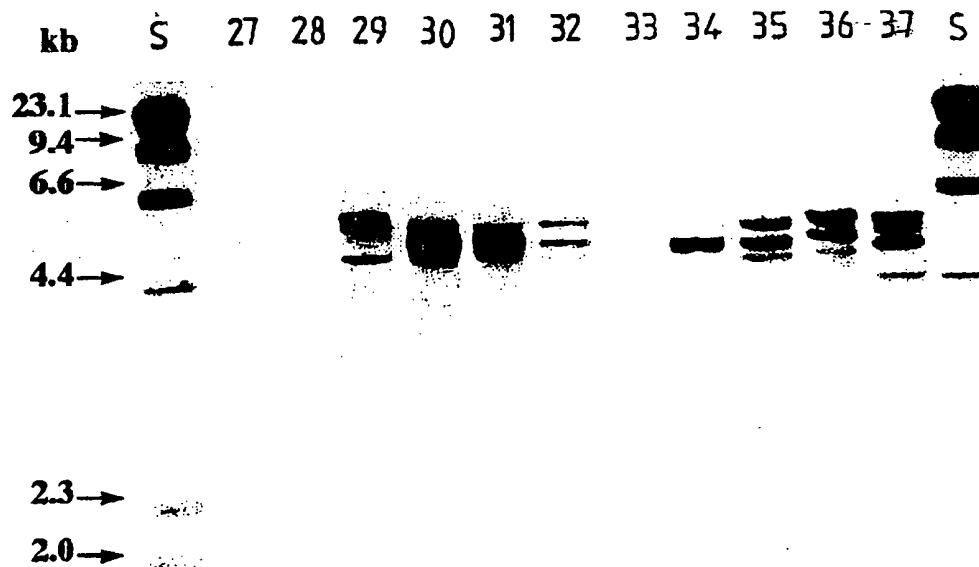


FIG.13

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 93 23568 A (A. R. HOLMES ET AL.) 25 November 1993 cited in the application see page 11, line 7 - page 12, line 35 see page 18, line 32 - page 19, line 27	1-6,10, 11,14
Y	---	7-9,12
X	WO 95 11991 A (UNIVERSITA CLAUDE BERNARD) 4 May 1995 cited in the application see the whole document	1-6,10, 11,14
Y	---	7-9,12
	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

1 July 1997

Date of mailing of the international search report

18. 07. 97

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. ( - 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax ( - 31-70) 340-3016

Authorized officer

De Kok, A

## INTERNATIONAL SEARCH REPORT

Intern. Patent Application No

PCT/FR 97/00518

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 62, no. 1, January 1996, WASHINGTON US, pages 41-46, XP000196695 M.M. BALEIRAS COUTO ET AL.: "Evaluation of molecular typing techniques to assign genetic diversity among Saccharomyces cerevisiae strains" cited in the application see the whole document ---	1
P,X	JOURNAL OF CLINICAL MICROBIOLOGY, vol. 34, no. 7, July 1996, WASHINGTON US, pages 1726-1731, XP000196696 A. CARLOTTI ET AL.: "Species-specific identification of Candida krusei by hybridization with the CkF1,2 DNA probe" see the whole document -----	1-14

Patent document cited in search report	Publication date -	Patent family member(s)	Publication date
WO 9323568 A	25-11-93	AU 4094293 A	13-12-93
		CA 2136206 A	25-11-93
		EP 0642588 A	15-03-95
-----			
WO 9511991 A	04-05-95	FR 2711669 A	05-05-95
		CA 2175227 A	04-05-95
		EP 0725836 A	14-08-96
-----			

**A. CLASSEMENT DE L'OBJET DE LA DEMANDE**  
 CIB 6 C12Q1/68

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

**B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE**

 Documentation minimale consultée (système de classification suivi des symboles de classement)  
 CIB 6 C12Q

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

**C. DOCUMENTS CONSIDERES COMME PERTINENTS**

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 93 23568 A (A. R. HOLMES ET AL.) 25 Novembre 1993 cité dans la demande voir page 11, ligne 7 - page 12, ligne 35 voir page 18, ligne 32 - page 19, ligne 27	1-6,10, 11,14
Y	---	7-9,12
X	WO 95 11991 A (UNIVERSITA CLAUDE BERNARD) 4 Mai 1995 cité dans la demande	1-6,10, 11,14
Y	voir le document en entier ---	7-9,12
	--- -/--	

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

\* Catégories spéciales de documents cités:

- \* "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- \* "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- \* "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- \* "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- \* "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent, l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent, l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&amp;" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

1 Juillet 1997

Date d'expédition du présent rapport de recherche internationale

18. 07. 97

Nom et adresse postale de l'administration chargée de la recherche internationale

 Office Européen des Brevets, P.B. 5818 Patentlaan 2  
 NL - 2280 HX Rijswijk  
 Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax (+ 31-70) 340-3016

Fonctionnaire autorisé

De Kok, A



## C(8a) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>APPLIED AND ENVIRONMENTAL MICROBIOLOGY, vol. 62, no. 1, Janvier 1996, WASHINGTON US, pages 41-46, XP000196695 M.M. BALEIRAS COUTO ET AL.: "Evaluation of molecular typing techniques to assign genetic diversity among Saccharomyces cerevisiae strains" cité dans la demande voir le document en entier</p>	1
P,X	<p>--- JOURNAL OF CLINICAL MICROBIOLOGY, vol. 34, no. 7, Juillet 1996, WASHINGTON US, pages 1726-1731, XP000196696 A. CARLOTTI ET AL.: "Species-specific identification of Candida krusei by hybridization with the CkF1,2 DNA probe" voir le document en entier</p> <p>-----</p>	1-14

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9323568 A	25-11-93	AU 4094293 A	13-12-93
		CA 2136206 A	25-11-93
		EP 0642588 A	15-03-95
-----			
WO 9511991 A	04-05-95	FR 2711669 A	05-05-95
		CA 2175227 A	04-05-95
		EP 0725836 A	14-08-96
-----			